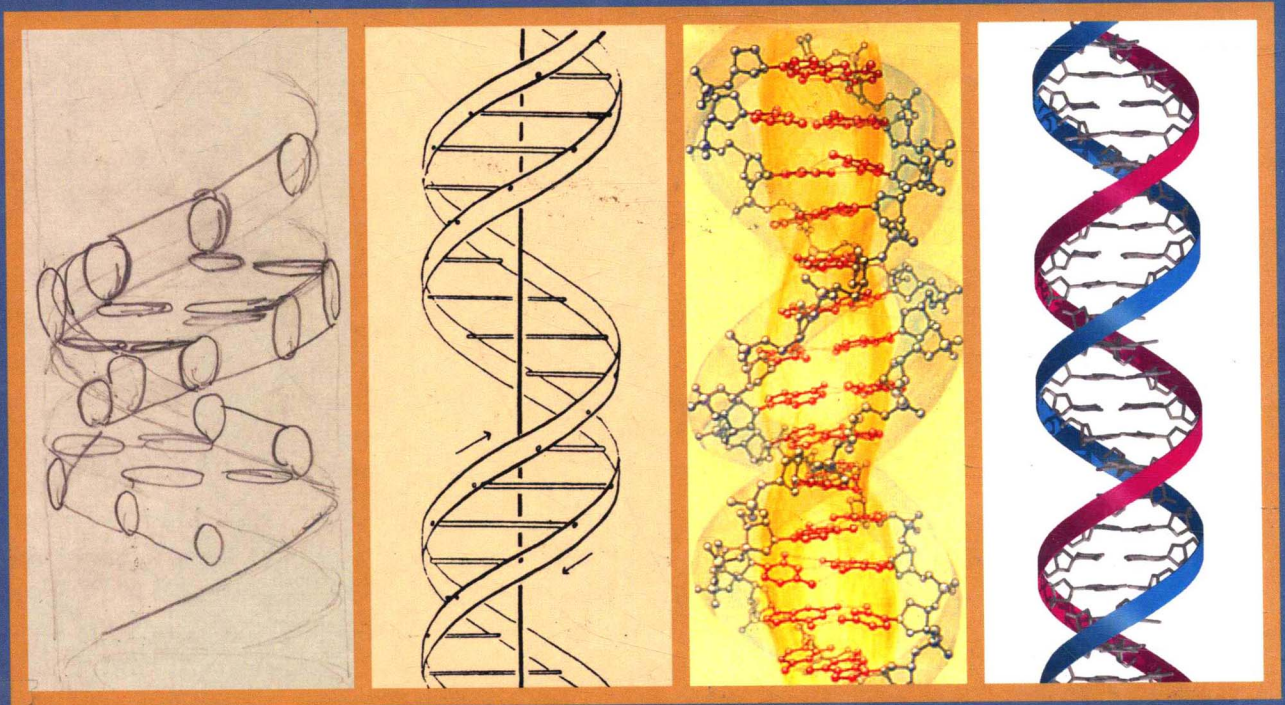


MOLECULAR BIOLOGY OF THE GENE

SEVENTH EDITION



WATSON • BAKER • BELL
GANN • LEVINE • LOSICK

MOLECULAR BIOLOGY^V OF THE GENE

SEVENTH EDITION

JAMES D. WATSON
Cold Spring Harbor Laboratory

ALEXANDER GANN
Cold Spring Harbor Laboratory

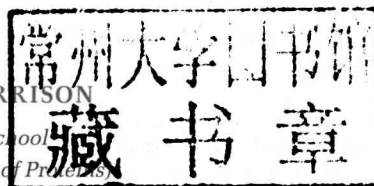
TANIA A. BAKER
Massachusetts Institute of Technology

MICHAEL LEVINE
University of California, Berkeley

STEPHEN P. BELL
Massachusetts Institute of Technology

RICHARD LOSICK
Harvard University

With
STEPHEN C. HARRISON
Harvard Medical School
(Chapter 6: The Structure of Proteins)



PEARSON

Boston Columbus Indianapolis New York San Francisco Upper Saddle River
Amsterdam Cape Town Dubai London Madrid Milan Munich Paris Montréal Toronto
Delhi Mexico City São Paulo Sydney Hong Kong Seoul Singapore Taipei Tokyo



COLD SPRING HARBOR LABORATORY PRESS
Cold Spring Harbor, New York

PEARSON

Editor-in-Chief: Beth Wilbur
Senior Acquisitions Editor: Josh Frost
Executive Director of Development: Deborah Gale
Assistant Editor: Katherine Harrison-Adcock
Managing Editor: Michael Early
Production Project Manager: Lori Newman
Illustrators: Dragonfly Media Group
Manufacturing Buyer: Michael Penne
Director of Marketing: Christy Lesko
Executive Marketing Manager: Lauren Harp
Executive Media Producer: Laura Tommasi
Editorial Media Producer: Lee Ann Doctor
Supervising Media Project Manager: David Chavez
Director of Content Development, MasteringBiology: Natania Mlawer
Content Specialist, MasteringBiology: J. Zane Barlow, PhD

COLD SPRING HARBOR LABORATORY PRESS

Publisher and Sponsoring Editor: John Inglis
Editorial Director: Alexander Gann
Director of Editorial Development: Jan Argentine
Managing Editor and Developmental Editor: Kaaren Janssen
Project Manager: Inez Sialiano
Production Manager: Denise Weiss
Production Editor: Kathleen Bubbeo
Permissions Coordinator: Carol Brown
Crystal Structure Images: Leemor Joshua-Tor and Stephen C. Harrison
Cover Designer: Mike Albano

Front and Back Cover Images: Far left, drawing by Francis Crick, Wellcome Library, London. Second from left, from Watson J.D. and Crick F.H.C. 1953. *Nature* **171**: 737–738. Second from right, Irving Geis illustration. Rights owned by Howard Hughes Medical Institute. Not to be reproduced without permission. Far right, structure by Leemor Joshua-Tor (image prepared with PyMOL).

Credits and acknowledgments for materials borrowed from other sources and reproduced, with permission, in this textbook appear on the appropriate page within the text.

Copyright © 2014, 2008, 2004 Pearson Education, Inc. All rights reserved. Manufactured in the United States of America. This publication is protected by Copyright, and permission should be obtained from the publisher prior to any prohibited reproduction, storage in a retrieval system, or transmission in any form or by any means, electronic, mechanical, photocopying, recording, or likewise. To obtain permission(s) to use material from this work, please submit a written request to Pearson Education, Inc., Permissions Department, 1900 E. Lake Ave., Glenview, IL 60025. For information regarding permissions, call (847) 486-2635.

Readers may view, browse, and/or download material for temporary copying purposes only, provided these uses are for noncommercial personal purposes. Except as provided by law, this material may not be further reproduced, distributed, transmitted, modified, adapted, performed, displayed, published, or sold in whole or in part, without prior written permission from the publisher.

Many of the designations used by manufacturers and sellers to distinguish their products are claimed as trademarks. Where those designations appear in this book, and the publisher was aware of a trademark claim, the designations have been printed in initial caps or all caps.

MasteringBiology and BioFlix are trademarks, in the U.S. and/or other countries, of Pearson Education, Inc. or its affiliates.

Library of Congress Cataloging-in-Publication Data

Watson, James D.

Molecular biology of the gene / James D. Watson, Cold Spring Harbor Laboratory, Tania A. Baker, Massachusetts Institute of Technology, Alexander Gann, Cold Spring Harbor Laboratory, Michael Levine, University of California, Berkeley, Richard Losick, Harvard University.
pages cm

Includes bibliographical references and index.

ISBN-13: 978-0-321-76243-6 (hardcover (student ed))

ISBN-10: 0-321-76243-6 (hardcover (student ed))

ISBN-13: 978-0-321-90537-6 (paper (a la carte))

ISBN-10: 0-321-90537-7 (paper (a la carte))

[etc.]

1. Molecular biology--Textbooks. 2. Molecular genetics--Textbooks. I. Title.

QH506.M6627 2013

572'.33--dc23

2012046093

2 3 4 5 6 7 8 9 10—DOW—17 16 15 14 13

PEARSON

www.pearsonhighered.com



COLD SPRING HARBOR
LABORATORY PRESS
www.cshlpress.org

ISBN 10: 0-321-76243-6 (Student Edition)

ISBN 13: 978-0-321-76243-6 (Student Edition)

ISBN 10: 0-321-90264-5 (Instructor's Review Copy)

ISBN 13: 978-0-321-90264-1 (Instructor's Review Copy)

ISBN 10: 0-321-90537-7 (Books à la Carte Edition)

ISBN 13: 978-0-321-90537-6 (Books à la Carte Edition)

Preface

THE NEW EDITION OF *MOLECULAR BIOLOGY OF THE GENE* appears in this, its 7th edition, on the 60th anniversary of the discovery of the structure of DNA in 1953, an occasion celebrated by our cover design. The double-helical structure, held together by specific pairing between the bases on the two strands, has become one of the iconic images of science. The image of the microscope was perhaps the icon of science in the late 19th century, displaced by the mid 20th century by the graphical representation of the atom with its orbiting electrons. But by the end of the century that image had in turn given way to the double helix.

The field of molecular biology as we understand it today was born out of the discovery of the DNA structure and the agenda for research that that structure immediately provided. The paper by Watson and Crick proposing the double helix ended with a now famous sentence: "It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material." The structure suggested how DNA could replicate, opening the way to investigate, in molecular terms, how genes are passed down through generations. It was also immediately apparent that the order of bases along a DNA molecule could represent a "genetic code," and so an attack on that second great mystery of genetics—how genes encode characteristics—could also be launched.

By the time the first edition of *Molecular Biology of the Gene* was published, just 12 years later in 1965, it had been confirmed that DNA replicated in the manner suggested by the model, the genetic code had all but been cracked, and the mechanism by which genes are expressed, and how that expression is regulated, had been established at least in outline. The field of molecular biology was ripe for its first textbook, defining for the first time the curriculum for undergraduate courses in this topic.

Our understanding of the mechanisms underlying these processes has hugely increased over the last 48 years since that first edition, often driven by technological advances, including DNA sequencing (another anniversary this year is the 10th anniversary of completion of the human genome project). The current edition of *Molecular Biology of the Gene* celebrates both the central intellectual framework of the field put in place in that first edition and the extraordinary mechanistic, biological, and evolutionary understanding that has since been achieved.

New to This Edition

There are a number of major changes to the new edition. As well as wide-ranging updates, these include changes in organization, addition of completely new chapters, and the addition of new topics within existing chapters.

- **New Part 2 on the Structure and Study of Macromolecules.** In this new section, each of the three major macromolecules gets its own chapter. The DNA chapter is retained from the previous edition, but what was previously just a short section at the end of that chapter is now expanded into a whole new chapter on the structure of RNA. The chapter on the structure of proteins is completely new and was written for this edition by Stephen Harrison (Harvard University).

- *Techniques chapter moved from the end of the book into Part 2.* This revised and relocated chapter introduces the important techniques that will be referred to throughout the book. In addition to many of the basic techniques of molecular biology, this chapter now includes an updated section on many genomics techniques routinely employed by molecular biologists. Techniques more specialized for particular chapters appear as boxes within those chapters.
- *Completely new chapter on The Origin and Early Evolution of Life.* This chapter shows how the techniques of molecular biology and biochemistry allow us to consider—even reconstruct—how life might have arisen and addresses the prospect of creating life in a test tube (synthetic biology). The chapter also reveals how, even at the very early stages of life, molecular processes were subject to evolution.
- *New material on many aspects of gene regulation.* Part 5 of the book is concerned with gene regulation. In this edition we have introduced significant new topics, such as quorum sensing in bacterial populations, the bacterial CRISPR defense system and piRNAs in animals, the function of Polycomb, and increased discussion of other so-called “epigenetic” mechanisms of gene regulation in higher eukaryotes. The regulation of “paused polymerase” at many genes during animal development and the critical involvement of nucleosome positioning and remodeling at promoters during gene activation are also new topics to this edition.
- *End-of-chapter questions.* Appearing for the first time in this edition, these include both short answer and data analysis questions. The answers to the even-numbered questions are included as Appendix 2 at the back of the book.
- *New experiments and experimental approaches reflecting recent advances in research.* Integrated within the text are new experimental approaches and applications that broaden the horizons of research. These include, for example, a description of how the genetic code can be experimentally expanded to generate novel proteins, creation of a synthetic genome to identify the minimal features required for life, discussion of new genome-wide analysis of nucleosome positioning, experiments on bimodal switches in bacteria, and how new antibacterial drugs are being designed that target the quorum-sensing pathways required for pathogenesis.

Supplements

MasteringBiology www.masteringbiology.com

MasteringBiology is an online homework, tutorial, and assessment system that delivers self-paced tutorials that provide individualized coaching, focus on your course objectives, and are responsive to each student's progress. The Mastering system helps instructors maximize class time with customizable, easy-to-assign, and automatically graded assessments that motivate students to learn outside of class and arrive prepared for lecture. MasteringBiology includes the book's end-of-chapter problems, eighteen 3D structure tutorials, reading quizzes, animations, videos, and a wide variety of activities. The eText is also available through MasteringBiology, providing access to the complete textbook and featuring powerful interactive and customization functions.

Instructor Resource DVD 978-0-321-88342-1/0-321-88342-X

Available free to all adopters, this dual-platform DVD-ROM contains all art and tables from the book in JPEG and PowerPoint in high-resolution (150 dpi) files. The PowerPoint slides include problems formatted for use with Classroom Response Systems. This DVD-ROM also contains an answer key for all of the end-of-chapter Critical Thinking questions included in MasteringBiology.

Transparency Acetates 978-0-321-88341-4/0-321-88341-1

Features approximately 90 four-color illustrations from the text. These transparencies are free to all adopters.

Cold Spring Harbor Laboratory Photographs

As in the previous edition, each part opener includes photographs, some newly added to this edition. These pictures, selected from the archives of Cold Spring Harbor Laboratory, were all taken at the Lab, the great majority during the Symposia hosted there almost every summer since 1933. Captions identify who is in each picture and when it was taken. Many more examples of these historic photos can be found at the CSHL archives website (<http://archives.cshl.edu/>).

Acknowledgments

Parts of the current edition grew out of an introductory course on molecular biology taught by one of us (RL) at Harvard University, and this author is grateful to Steve Harrison and Jim Wang who contributed to this course in past years. In the case of Steve Harrison, we are additionally indebted to him for writing and illustrating a brand new chapter on protein structure especially for this new edition. No one could be better qualified for such a task, and we are the grateful beneficiaries of—and the book is immeasurably improved by—his contribution.

We are also grateful to Craig Hunter, who earlier wrote the section on the worm for Appendix 1, and to Rob Martienssen, who wrote the section on plants for that same appendix.

We have shown sections of the manuscript to various colleagues and their comments have been extremely helpful. Specifically we thank Katsura Asano, Stephen Blacklow, Jamie Cate, Amy Caudy, Irene Chen, Victoria D'Souza, Richard Ebright, Mike Eisen, Chris Fromme, Brenton Graveley, Chris Hammell, Steve Hahn, Oliver Hobert, Ann Hochschild, Jim Hu, David Jerulzalmi, Leemor Joshua-Tor, Sandy Johnson, Andrew Knoll, Adrian Krainer, Julian Lewis, Sue Lovett, Karolin Luger, Kristen Lynch, Rob Martienssen, Bill McGinnis, Matt Michael, Lily Mirels, Nipam Patel, Mark Ptashne, Danny Reinberg, Dimitar Sasselov, David Shechner, Sarah T. Stewart-Mukhopadhyay, Bruce Stillman, and Jack Szostak.

We also thank those who provided us with figures, or the wherewithal to create them: Sean Carroll, Seth Darst, Paul Fransz, Brenton Graveley, Ann Hochschild, Julian Lewis, Bill McGinnis, Phoebe Rice, Dan Rokhsar, Nori Satoh, Matt Scott, Ali Shilatifard, Peter Sorger, Tom Steitz, Andrzej Stasiak, Dan Voytas, and Steve West.

New to this edition are end-of-chapter questions, provided by Mary Ellen Wiltrout, and we thank her for these efforts that have enhanced the new edition. In addition, Mary Ellen helped with revisions to the DNA repair chapter.

We are indebted to Leemor Joshua-Tor, who so beautifully rendered the majority of the structure figures throughout the book. Her skill and patience are much appreciated.

We are also grateful to those who provided their software¹: Per Kraulis, Robert Esnouf, Ethan Merritt, Barry Honig, and Warren Delano. Coordinates were obtained from the Protein Data Bank (www.rcsb.org/pdb/), and citations to those who solved each structure are included in the figure legends.

Our art program was again executed by a team from the Dragonfly Media Group, led by Craig Durant. Denise Weiss and Mike Albano produced a beautiful cover design. We thank Clare Bunce and the CSHL Archive for providing the photos for the part openers and for much help tracking them down.

We thank Josh Frost at Pearson who oversaw our efforts and was always on hand to help us out or provide advice. In development at CSHL Press, Jan Argentine provided great support, guidance, and perspective throughout the process. Our heartfelt thanks to Kaaren Janssen who was once again our constant savior—editing and organizing, encouraging and understanding—and unstintingly good-humored even on the darkest days. Inez Sialiano kept track of the output, and Carol Brown dealt with the permissions as efficiently as ever. In production, we relied heavily on the extraordinary efforts and patience

of Kathleen Bubbeo, for which we are most grateful. And we must also thank Denise Weiss, who oversaw production and ensured that the book looked so good by finessing the page layout and creating the design. John Inglis as ever created the environment in which this could all take place.

And once again, we thank our families for putting up with this book for a third time!

JAMES D. WATSON
TANIA A. BAKER
STEPHEN P. BELL
ALEXANDER GANN
MICHAEL LEVINE
RICHARD LOSICK

¹Per Kraulis granted permission to use MolScript (Kraulis P.J. 1991. MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. *J. Appl. Cryst.* **24**: 946–950). Robert Esnouf gave permission to use BobScript (Esnouf R.M. 1997. *J. Mol. Graph.* **15**: 132–134). In addition, Ethan Merritt gave us use of Raster3D (Merritt E.A. and Bacon D.J. 1997. Raster3D: Photorealistic molecular graphics. *Methods Enzymol.* **277**: 505–524), and Barry Honig granted permission to use GRASP (Nicolls A., Sharp K.A., and Honig B. 1991. Protein folding and association: Insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins* **11**: 281–296). Warren DeLano agreed to the use of PyMOL (DeLano W.L. 2002. *The PyMOL Molecular Graphics System*. DeLano Scientific, Palo Alto, California).

About the Authors

JAMES D. WATSON is Chancellor Emeritus at Cold Spring Harbor Laboratory, where he was previously its Director from 1968 to 1993, President from 1994 to 2003, and Chancellor from 2003 to 2007. He spent his undergraduate years at the University of Chicago and received his Ph.D. in 1950 from Indiana University. Between 1950 and 1953, he did post-doctoral research in Copenhagen and Cambridge, England. While at Cambridge, he began the collaboration that resulted in the elucidation of the double-helical structure of DNA in 1953. (For this discovery, Watson, Francis Crick, and Maurice Wilkins were awarded the Nobel Prize in 1962.) Later in 1953, he went to the California Institute of Technology. He moved to Harvard in 1955, where he taught and did research on RNA synthesis and protein synthesis until 1976. He was the first Director of the National Center for Genome Research of the National Institutes of Health from 1989 to 1992. Dr. Watson was sole author of the first, second, and third editions of *Molecular Biology of the Gene*, and a co-author of the fourth, fifth and sixth editions. These were published in 1965, 1970, 1976, 1987, 2003, and 2007, respectively. He is also a co-author of two other textbooks, *Molecular Biology of the Cell* and *Recombinant DNA*, as well as author of the celebrated 1968 memoir, *The Double Helix*, which in 2012 was listed by the Library of Congress as one of the 88 Books That Shaped America.

TANIA A. BAKER is the Head of the Department and Whitehead Professor of Biology at the Massachusetts Institute of Technology and an Investigator of the Howard Hughes Medical Institute. She received a B.S. in biochemistry from the University of Wisconsin, Madison, and a Ph.D. in biochemistry from Stanford University in 1988. Her graduate research was carried out in the laboratory of Professor Arthur Kornberg and focused on mechanisms of initiation of DNA replication. She did postdoctoral research in the laboratory of Dr. Kiyoshi Mizuuchi at the National Institutes of Health, studying the mechanism and regulation of DNA transposition. Her current research explores mechanisms and regulation of genetic recombination, enzyme-catalyzed protein unfolding, and ATP-dependent protein degradation. Professor Baker received the 2001 Eli Lilly Research Award from the American Society of Microbiology and the 2000 MIT School of Science Teaching Prize for Undergraduate Education and is a Fellow of the American Academy of Arts and Sciences since 2004 and was elected to the National Academy of Sciences in 2007. She is co-author (with Arthur Kornberg) of the book *DNA Replication*, Second Edition.

STEPHEN P. BELL is a Professor of Biology at the Massachusetts Institute of Technology and an Investigator of the Howard Hughes Medical Institute. He received B.A. degrees from the Department of Biochemistry, Molecular Biology, and Cell Biology and the Integrated Sciences Program at Northwestern University and a Ph.D. in biochemistry at the University of California, Berkeley, in 1991. His graduate research was carried out in the laboratory of Dr. Robert Tjian and focused on eukaryotic transcription. He did postdoctoral research in the laboratory of Dr. Bruce Stillman at Cold Spring Harbor Laboratory, working on the initiation of eukaryotic DNA replication. His current research focuses on the mechanisms controlling the duplication of eukaryotic chromosomes. Professor Bell received the 2001 ASBMB–Schering Plough Scientific Achievement Award, the

1998 Everett Moore Baker Memorial Award for Excellence in Undergraduate Teaching at MIT, the 2006 MIT School of Science Teaching Award, and the 2009 National Academy of Sciences Molecular Biology Award.

ALEXANDER GANN is the Lita Annenberg Hazen Dean and Professor in the Watson School of Biological Sciences at Cold Spring Harbor Laboratory. He is also a Senior Editor at Cold Spring Harbor Laboratory Press. He received his B.Sc. in microbiology from University College London and a Ph.D. in molecular biology from The University of Edinburgh in 1989. His graduate research was carried out in the laboratory of Noreen Murray and focused on DNA recognition by restriction enzymes. He did postdoctoral research in the laboratory of Mark Ptashne at Harvard, working on transcriptional regulation, and that of Jeremy Brockes at the Ludwig Institute of Cancer Research at University College London, where he worked on newt limb regeneration. He was a Lecturer at Lancaster University, United Kingdom, from 1996 to 1999, before moving to Cold Spring Harbor Laboratory. He is co-author (with Mark Ptashne) of the book *Genes & Signals* (2002) and co-editor (with Jan Witkowski) of *The Annotated and Illustrated Double Helix* (2012).

MICHAEL LEVINE is a Professor of Genetics, Genomics and Development at the University of California, Berkeley, and is also Co-Director of the Center for Integrative Genomics. He received his B.A. from the Department of Genetics at the University of California, Berkeley, and his Ph.D. with Alan Garen in the Department of Molecular Biophysics and Biochemistry from Yale University in 1981. As a Postdoctoral Fellow with Walter Gehring and Gerry Rubin from 1982 to 1984, he studied the molecular genetics of *Drosophila* development. Professor Levine's research group currently studies the gene networks responsible for the gastrulation of the *Drosophila* and *Ciona* (sea squirt) embryos. He holds the F. Williams Chair in Genetics and Development at University of California, Berkeley. He was awarded the Monsanto Prize in Molecular Biology from the National Academy of Sciences in 1996 and was elected to the American Academy of Arts and Sciences in 1996 and the National Academy of Sciences in 1998.

RICHARD LOSICK is the Maria Moors Cabot Professor of Biology, a Harvard College Professor, and a Howard Hughes Medical Institute Professor in the Faculty of Arts and Sciences at Harvard University. He received his A.B. in chemistry at Princeton University and his Ph.D. in biochemistry at the Massachusetts Institute of Technology. Upon completion of his graduate work, Professor Losick was named a Junior Fellow of the Harvard Society of Fellows when he began his studies on RNA polymerase and the regulation of gene transcription in bacteria. Professor Losick is a past Chairman of the Departments of Cellular and Developmental Biology and Molecular and Cellular Biology at Harvard University. He received the Camille and Henry Dreyfus Teacher-Scholar Award and is a member of the National Academy of Sciences, a Fellow of the American Academy of Arts and Sciences, a Fellow of the American Association for the Advancement of Science, a Fellow of the American Academy of Microbiology, a member of the American Philosophical Society, and a former Visiting Scholar of the Phi Beta Kappa Society. Professor Losick is the 2007 winner of the Selman A. Waksman Award of the National Academy of Sciences, a 2009 winner of the Canada Gairdner Award, a 2012 winner of the Louisa Gross Horwitz Prize for Biology or Biochemistry of Columbia University, and a 2012 winner of the Harvard University Fannie Cox Award for Excellence in Science Teaching.

Class Testers and Reviewers

We wish to thank all of the instructors for their thoughtful suggestions and comments on versions of many chapters in this book.

Chapter Reviewers

Ann Aguanno, *Marymount Manhattan College*

David P. Aiello, *Austin College*

Charles F. Austerberry, *Creighton University*

David G. Bear, *University of New Mexico Health Sciences Center*

Margaret E. Beard, *College of the Holy Cross*

Gail S. Begley, *Northeastern University*

Sanford Bernstein, *San Diego State University*

Michael Blaber, *Florida State University*

Nicole Bournias, *California State University, San Bernardino*

John Boyle, *Mississippi State University*

Suzanne Bradshaw, *University of Cincinnati*

John G. Burr, *University of Texas at Dallas*

Michael A. Campbell, *Pennsylvania State University, Erie, The Behrend College*

Aaron Cassill, *University of Texas at San Antonio*

Shirley Coomber, *King's College, University of London*

Anne Cordon, *University of Toronto*

Sumana Datta, *Texas A&M University*

Jeff DeJong, *University of Texas at Dallas*

Jurgen Denecke, *University of Leeds*

Susan M. DiBartolomeis, *Millersville University*

Santosh R. D'Mello, *University of Texas at Dallas*

Robert J. Duronio, *University of North Carolina, Chapel Hill*

Steven W. Edwards, *University of Liverpool*

David Frick, *University of Wisconsin*

Allen Gathman, *Southeast Missouri State University*

Anthony D.M. Glass, *University of British Columbia*

Elliott S. Goldstein, *Arizona State University*

Ann Grens, *Indiana University, South Bend*

Gregory B. Hecht, *Rowan University*

Robert B. Helling, *University of Michigan*

David C. Higgs, *University of Wisconsin, Parkside*

Mark Kainz, *Colgate University*

Gregory M. Kelly, *University of Western Ontario*

Ann Kleinschmidt, *Allegheny College*

Dan Krane, *Wright State University*

Mark Levinthal, *Purdue University*

Gary J. Lindquister, *Rhodes College*

James Lodolce, *Loyola University Chicago*

Curtis Loer, *University of San Diego*

Virginia McDonough, *Hope College*

Michael J. McPherson, *University of Leeds*

Victoria Meller, *Tufts University*

William L. Miller, *North Carolina State University*

Dragana Miskovic, *University of Waterloo*

David Mullin, *Tulane University*

Jeffrey D. Newman, *Lycoming College*

James B. Olesen, *Ball State University*

Anthony J. Otsuka, *Illinois State University*

Karen Palter, *Temple University*

James G. Patton, *Vanderbilt University*

Ian R. Phillips, *Queen Mary, University of London*

Steve Picksley, *University of Bradford*

Debra Pires, *University of California, Los Angeles*

Todd P. Primm, *University of Texas at El Paso*

Phillip E. Ryals, *The University of West Florida*

Eva Sapi, *University of New Haven*

Jon B. Scales, *Midwestern State University*

Michael Schultze, *University of York*

Venkat Sharma, *University of West Florida*

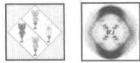
Erica L. Shelley, *University of Toronto at Mississauga*
Elizabeth A. Shephard, *University College, London*
Margaret E. Stevens, *Ripon College*
Akif Uzman, *University of Houston, Downtown*
Quinn Vega, *Montclair State University*
Jeffrey M. Voight, *Albany College of Pharmacy*
Lori L. Wallrath, *University of Iowa*
Robert Wiggers, *Stephen F. Austin State University*
Bruce C. Wightman, *Muhlenberg College*
Bob Zimmermann, *University of Massachusetts*

Class Testers

Charles F. Austerberry, *Creighton University*
Christine E. Bezotté, *Elmira College*
Astrid Helfant, *Hamilton College*
Gerald Joyce, *The Scripps Research Institute*
Jocelyn Krebs, *University of Alaska, Anchorage*
Cran Lucas, *Louisiana State University in Shreveport*
Anthony J. Otsuka, *Illinois State University*
Charles Polson, *Florida Institute of Technology*
Ming-Che Shih, *University of Iowa*

Brief Contents

PART 1



HISTORY, 1

- 1 The Mendelian View of the World, 5
- 2 Nucleic Acids Convey Genetic Information, 21

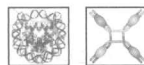
PART 2



STRUCTURE AND STUDY OF MACROMOLECULES, 45

- 3 The Importance of Weak and Strong Chemical Bonds, 51
- 4 The Structure of DNA, 77
- 5 The Structure and Versatility of RNA, 107
- 6 The Structure of Proteins, 121
- 7 Techniques of Molecular Biology, 147

PART 3



MAINTENANCE OF THE GENOME, 193

- 8 Genome Structure, Chromatin, and the Nucleosome, 199
- 9 The Replication of DNA, 257
- 10 The Mutability and Repair of DNA, 313
- 11 Homologous Recombination at the Molecular Level, 341
- 12 Site-Specific Recombination and Transposition of DNA, 377

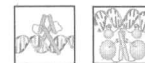
PART 4



EXPRESSION OF THE GENOME, 423

- 13 Mechanisms of Transcription, 429
- 14 RNA Splicing, 467
- 15 Translation, 509
- 16 The Genetic Code, 573
- 17 The Origin and Early Evolution of Life, 593

PART 5



REGULATION, 609

- 18 Transcriptional Regulation in Prokaryotes, 615
- 19 Transcriptional Regulation in Eukaryotes, 657
- 20 Regulatory RNAs, 701
- 21 Gene Regulation in Development and Evolution, 733
- 22 Systems Biology, 775

PART 6



APPENDICES, 793

- 1 Model Organisms, 797
- 2 Answers, 831

Index, 845

PART 2: STRUCTURE AND STUDY OF MACROMOLECULES, 45



3 The Importance of Weak and Strong Chemical Bonds, 51

CHARACTERISTICS OF CHEMICAL BONDS, 51

- Chemical Bonds Are Explainable in Quantum-Mechanical Terms, 52
- Chemical-Bond Formation Involves a Change in the Form of Energy, 53
- Equilibrium between Bond Making and Breaking, 53

THE CONCEPT OF FREE ENERGY, 54

- K_{eq} Is Exponentially Related to ΔG , 54
- Covalent Bonds Are Very Strong, 54

WEAK BONDS IN BIOLOGICAL SYSTEMS, 55

- Weak Bonds Have Energies between 1 and 7 kcal/mol, 55
- Weak Bonds Are Constantly Made and Broken at Physiological Temperatures, 55
- The Distinction between Polar and Nonpolar Molecules, 55
- van der Waals Forces, 56
- Hydrogen Bonds, 57
- Some Ionic Bonds Are Hydrogen Bonds, 58
- Weak Interactions Demand Complementary Molecular Surfaces, 58
- Water Molecules Form Hydrogen Bonds, 59
- Weak Bonds between Molecules in Aqueous Solutions, 59
- Organic Molecules That Tend to Form Hydrogen Bonds Are Water Soluble, 60
- Hydrophobic "Bonds" Stabilize Macromolecules, 60
- ADVANCED CONCEPTS BOX 3-1 The Uniqueness of Molecular Shapes and the Concept of Selective Stickiness, 61*
- The Advantage of ΔG between 2 and 5 kcal/mol, 62

- Weak Bonds Attach Enzymes to Substrates, 62
- Weak Bonds Mediate Most Protein–DNA and Protein–Protein Interactions, 62

HIGH-ENERGY BONDS, 63

MOLECULES THAT DONATE ENERGY ARE THERMODYNAMICALLY UNSTABLE, 63

ENZYMES LOWER ACTIVATION ENERGIES IN BIOCHEMICAL REACTIONS, 65

FREE ENERGY IN BIOMOLECULES, 66

- High-Energy Bonds Hydrolyze with Large Negative ΔG , 66

HIGH-ENERGY BONDS IN BIOSYNTHETIC REACTIONS, 67

- Peptide Bonds Hydrolyze Spontaneously, 68
- Coupling of Negative with Positive ΔG , 69

ACTIVATION OF PRECURSORS IN GROUP TRANSFER REACTIONS, 69

- ATP Versatility in Group Transfer, 70
- Activation of Amino Acids by Attachment of AMP, 70
- Nucleic Acid Precursors Are Activated by the Presence of $P \sim P$, 71
- The Value of $P \sim P$ Release in Nucleic Acid Synthesis, 72
- $P \sim P$ Splits Characterize Most Biosynthetic Reactions, 73

SUMMARY, 74

BIBLIOGRAPHY, 75

QUESTIONS, 75



4 The Structure of DNA, 77

DNA STRUCTURE, 78

- DNA Is Composed of Polynucleotide Chains, 78
- Each Base Has Its Preferred Tautomeric Form, 80
- The Two Strands of the Double Helix Are Wound around Each Other in an Antiparallel Orientation, 81
- The Two Chains of the Double Helix Have Complementary Sequences, 81

- The Double Helix Is Stabilized by Base Pairing and Base Stacking, 82
- Hydrogen Bonding Is Important for the Specificity of Base Pairing, 83
- Bases Can Flip Out from the Double Helix, 83
- DNA Is Usually a Right-Handed Double Helix, 83
- KEY EXPERIMENTS BOX 4-1 DNA Has 10.5 bp per Turn of the Helix in Solution: The Mica Experiment, 84*

The Double Helix Has Minor and Major Grooves, 84

The Major Groove Is Rich in Chemical Information, 85

The Double Helix Exists in Multiple Conformations, 86

DNA Can Sometimes Form a Left-Handed Helix, 87

KEY EXPERIMENTS BOX 4-2 How Spots on an X-Ray Film Reveal the Structure of DNA, 88

DNA Strands Can Separate (Denature) and Reassociate, 89

Some DNA Molecules Are Circles, 92

DNA TOPOLOGY, 93

Linking Number Is an Invariant Topological Property of Covalently Closed, Circular DNA, 93

Linking Number Is Composed of Twist and Writhe, 93

Lk^0 Is the Linking Number of Fully Relaxed cccDNA under Physiological Conditions, 94

DNA in Cells Is Negatively Supercoiled, 95

Nucleosomes Introduce Negative Supercoiling in Eukaryotes, 96

Topoisomerases Can Relax Supercoiled DNA, 97

Prokaryotes Have a Special Topoisomerase That Introduces Supercoils into DNA, 97

Topoisomerases Also Unknot and Disentangle DNA Molecules, 98

Topoisomerases Use a Covalent Protein–DNA Linkage to Cleave and Rejoin DNA Strands, 99

Topoisomerases Form an Enzyme Bridge and Pass DNA Segments through Each Other, 100

DNA Topoisomers Can Be Separated by Electrophoresis, 102

Ethidium Ions Cause DNA to Unwind, 102

KEY EXPERIMENTS BOX 4-3 Proving that DNA Has a Helical Periodicity of ~ 10.5 bp per Turn from the Topological Properties of DNA Rings, 103

SUMMARY, 103

BIBLIOGRAPHY, 104

QUESTIONS, 104



5 The Structure and Versatility of RNA, 107

RNA CONTAINS RIBOSE AND URACIL AND IS USUALLY SINGLE-STRANDED, 107

RNA CHAINS FOLD BACK ON THEMSELVES TO FORM LOCAL REGIONS OF DOUBLE HELIX SIMILAR TO A-FORM DNA, 108

RNA CAN FOLD UP INTO COMPLEX TERTIARY STRUCTURES, 110

NUCLEOTIDE SUBSTITUTIONS IN COMBINATION WITH CHEMICAL PROBING PREDICT RNA STRUCTURE, 111

MEDICAL CONNECTIONS BOX 5-1 An RNA Switch Controls Protein Synthesis by Murine Leukemia Virus, 112

DIRECTED EVOLUTION SELECTS RNAs THAT BIND SMALL MOLECULES, 114

SOME RNAs ARE ENZYMES, 114

TECHNIQUES BOX 5-2 Creating an RNA Mimetic of the Green Fluorescent Protein by Directed Evolution, 115

The Hammerhead Ribozyme Cleaves RNA by the Formation of a 2', 3' Cyclic Phosphate, 116

A Ribozyme at the Heart of the Ribosome Acts on a Carbon Center, 118

SUMMARY, 118

BIBLIOGRAPHY, 118

QUESTIONS, 118



6 The Structure of Proteins, 121

THE BASICS, 121

Amino Acids, 121

The Peptide Bond, 122

Polypeptide Chains, 123

Three Amino Acids with Special Conformational Properties, 124

ADVANCED CONCEPT BOX 6-1 Ramachandran Plot: Permitted Combinations of Main-Chain Torsion Angles ϕ and ψ , 124

IMPORTANCE OF WATER, 125

PROTEIN STRUCTURE CAN BE DESCRIBED AT FOUR LEVELS, 126

PROTEIN DOMAINS, 130

Polypeptide Chains Typically Fold into One or More Domains, 130

ADVANCED CONCEPTS BOX 6-2 Glossary of Terms, 130

Basic Lessons from the Study of Protein Structures, 131

Classes of Protein Domains, 132

Linkers and Hinges, 133

Post-Translational Modifications, 133

ADVANCED CONCEPTS BOX 6-3 The Antibody Molecule as an Illustration of Protein Domains, 133

FROM AMINO-ACID SEQUENCE TO THREE-DIMENSIONAL STRUCTURE, 134

Protein Folding, 134

KEY EXPERIMENTS BOX 6-4 Three-Dimensional Structure of a Protein Is Specified by Its Amino Acid Sequence (Anfinsen Experiment), 135

Predicting Protein Structure from Amino Acid Sequence, 135

CONFORMATIONAL CHANGES IN PROTEINS, 136



7 Techniques of Molecular Biology, 147

NUCLEIC ACIDS: BASIC METHODS, 148

Gel Electrophoresis Separates DNA and RNA Molecules according to Size, 148

Restriction Endonucleases Cleave DNA Molecules at Particular Sites, 149

DNA Hybridization Can Be Used to Identify Specific DNA Molecules, 151

Hybridization Probes Can Identify Electrophoretically Separated DNAs and RNAs, 151

Isolation of Specific Segments of DNA, 153

DNA Cloning, 154

Vector DNA Can Be Introduced into Host Organisms by Transformation, 155

Libraries of DNA Molecules Can Be Created by Cloning, 156

Hybridization Can Be Used to Identify a Specific Clone in a DNA Library, 156

Chemical Synthesis of Defined DNA Sequences, 157

The Polymerase Chain Reaction Amplifies DNAs by Repeated Rounds of DNA Replication In Vitro, 158

Nested Sets of DNA Fragments Reveal Nucleotide Sequences, 158

TECHNIQUES BOX 7-1 Forensics and the Polymerase Chain Reaction, 160

Shotgun Sequencing a Bacterial Genome, 162

The Shotgun Strategy Permits a Partial Assembly of Large Genome Sequences, 162

KEY EXPERIMENTS BOX 7-2 Sequenators Are Used for High-Throughput Sequencing, 163

The Paired-End Strategy Permits the Assembly of Large-Genome Scaffolds, 165

The \$1000 Human Genome Is within Reach, 167

PROTEINS AS AGENTS OF SPECIFIC MOLECULAR RECOGNITION, 137

Proteins That Recognize DNA Sequence, 137

Protein-Protein Interfaces, 140

Proteins That Recognize RNA, 141

ENZYMES: PROTEINS AS CATALYSTS, 141

REGULATION OF PROTEIN ACTIVITY, 142

SUMMARY, 143

BIBLIOGRAPHY, 144

QUESTIONS, 144

GENOMICS, 168

Bioinformatics Tools Facilitate the Genome-Wide Identification of Protein-Coding Genes, 169

Whole-Genome Tiling Arrays Are Used to Visualize the Transcriptome, 169

Regulatory DNA Sequences Can Be Identified by Using Specialized Alignment Tools, 171

Genome Editing Is Used to Precisely Alter Complex Genomes, 172

PROTEINS, 173

Specific Proteins Can Be Purified from Cell Extracts, 173

Purification of a Protein Requires a Specific Assay, 173

Preparation of Cell Extracts Containing Active Proteins, 174

Proteins Can Be Separated from One Another Using Column Chromatography, 174

Separation of Proteins on Polyacrylamide Gels, 176

Antibodies Are Used to Visualize Electrophoretically Separated Proteins, 176

Protein Molecules Can Be Directly Sequenced, 177

PROTEOMICS, 179

Combining Liquid Chromatography with Mass Spectrometry Identifies Individual Proteins within a Complex Extract, 179

Proteome Comparisons Identify Important Differences between Cells, 181

Mass Spectrometry Can Also Monitor Protein Modification States, 181

Protein-Protein Interactions Can Yield Information regarding Protein Function, 182

NUCLEIC ACID–PROTEIN INTERACTIONS, 182

The Electrophoretic Mobility of DNA Is Altered by Protein Binding, 183

DNA-Bound Protein Protects the DNA from Nucleases and Chemical Modification, 184

Chromatin Immunoprecipitation Can Detect Protein Association with DNA in the Cell, 185

Chromosome Conformation Capture Assays Are Used to Analyze Long-Range Interactions, 187

In Vitro Selection Can Be Used to Identify a Protein's DNA- or RNA-Binding Site, 189

BIBLIOGRAPHY, 190**QUESTIONS, 190**

PART 3: MAINTENANCE OF THE GENOME, 193

8 Genome Structure, Chromatin, and the Nucleosome, 199

GENOME SEQUENCE AND CHROMOSOME DIVERSITY, 200

Chromosomes Can Be Circular or Linear, 200

Every Cell Maintains a Characteristic Number of Chromosomes, 201

Genome Size Is Related to the Complexity of the Organism, 202

The *E. coli* Genome Is Composed Almost Entirely of Genes, 203

More Complex Organisms Have Decreased Gene Density, 204

Genes Make Up Only a Small Proportion of the Eukaryotic Chromosomal DNA, 205

The Majority of Human Intergenic Sequences Are Composed of Repetitive DNA, 207

CHROMOSOME DUPLICATION AND SEGREGATION, 208

Eukaryotic Chromosomes Require Centromeres, Telomeres, and Origins of Replication to Be Maintained during Cell Division, 208

Eukaryotic Chromosome Duplication and Segregation Occur in Separate Phases of the Cell Cycle, 210

Chromosome Structure Changes as Eukaryotic Cells Divide, 212

Sister-Chromatid Cohesion and Chromosome Condensation Are Mediated by SMC Proteins, 214

Mitosis Maintains the Parental Chromosome Number, 214

During Gap Phases, Cells Prepare for the Next Cell Cycle Stage and Check That the Previous Stage Is Completed Correctly, 217

Meiosis Reduces the Parental Chromosome Number, 217

Different Levels of Chromosome Structure Can Be Observed by Microscopy, 219

THE NUCLEOSOME, 220

Nucleosomes Are the Building Blocks of Chromosomes, 220

Histones Are Small, Positively Charged Proteins, 221

The Atomic Structure of the Nucleosome, 224

Histones Bind Characteristic Regions of DNA within the Nucleosome, 224

KEY EXPERIMENTS BOX 8-1 Micrococcal Nuclease and the DNA Associated with the Nucleosome, 226

Many DNA Sequence–Independent Contacts Mediate the Interaction between the Core Histones and DNA, 227

The Histone Amino-Terminal Tails Stabilize DNA Wrapping around the Octamer, 227

Wrapping of the DNA around the Histone Protein Core Stores Negative Superhelicity, 228

HIGHER-ORDER CHROMATIN STRUCTURE, 229

Heterochromatin and Euchromatin, 229

KEY EXPERIMENTS BOX 8-2 Nucleosomes and Superhelical Density, 230

Histone H1 Binds to the Linker DNA between Nucleosomes, 232

Nucleosome Arrays Can Form More Complex Structures: The 30-nm Fiber, 232

The Histone Amino-Terminal Tails Are Required for the Formation of the 30-nm Fiber, 234

Further Compaction of DNA Involves Large Loops of Nucleosomal DNA, 234

Histone Variants Alter Nucleosome Function, 234

REGULATION OF CHROMATIN STRUCTURE, 236

The Interaction of DNA with the Histone Octamer Is Dynamic, 236

Nucleosome-Remodeling Complexes Facilitate Nucleosome Movement, 237

Some Nucleosomes Are Found in Specific Positions: Nucleosome Positioning, 240

The Amino-Terminal Tails of the Histones Are Frequently Modified, 241

Protein Domains in Nucleosome-Remodeling and -Modifying Complexes Recognize Modified Histones, 244

KEY EXPERIMENTS BOX 8-3 *Determining Nucleosome Position in the Cell*, 245

Specific Enzymes Are Responsible for Histone Modification, 248

Nucleosome Modification and Remodeling Work Together to Increase DNA Accessibility, 249



9 The Replication of DNA, 257

THE CHEMISTRY OF DNA SYNTHESIS, 258

DNA Synthesis Requires Deoxynucleoside Triphosphates and a Primer:Template Junction, 258

DNA Is Synthesized by Extending the 3' End of the Primer, 259

Hydrolysis of Pyrophosphate Is the Driving Force for DNA Synthesis, 260

THE MECHANISM OF DNA POLYMERASE, 260

DNA Polymerases Use a Single Active Site to Catalyze DNA Synthesis, 260

TECHNIQUES BOX 9-1 *Incorporation Assays Can Be Used to Measure Nucleic Acid and Protein Synthesis*, 261

DNA Polymerases Resemble a Hand That Grips the Primer:Template Junction, 263

DNA Polymerases Are Processive Enzymes, 265

Exonucleases Proofread Newly Synthesized DNA, 267

MEDICAL CONNECTIONS BOX 9-2 *Anticancer and Antiviral Agents Target DNA Replication*, 268

THE REPLICATION FORK, 269

Both Strands of DNA Are Synthesized Together at the Replication Fork, 269

The Initiation of a New Strand of DNA Requires an RNA Primer, 270

RNA Primers Must Be Removed to Complete DNA Replication, 271

DNA Helicases Unwind the Double Helix in Advance of the Replication Fork, 272

DNA Helicase Pulls Single-Stranded DNA through a Central Protein Pore, 273

Single-Stranded DNA-Binding Proteins Stabilize ssDNA before Replication, 273

Topoisomerases Remove Supercoils Produced by DNA Unwinding at the Replication Fork, 275

Replication Fork Enzymes Extend the Range of DNA Polymerase Substrates, 275

NUCLEOSOME ASSEMBLY, 249

Nucleosomes Are Assembled Immediately after DNA Replication, 249

Assembly of Nucleosomes Requires Histone "Chaperones", 253

SUMMARY, 254

BIBLIOGRAPHY, 255

QUESTIONS, 255

THE SPECIALIZATION OF DNA POLYMERASES, 277

DNA Polymerases Are Specialized for Different Roles in the Cell, 277

Sliding Clamps Dramatically Increase DNA Polymerase Processivity, 278

Sliding Clamps Are Opened and Placed on DNA by Clamp Loaders, 281

ADVANCED CONCEPTS BOX 9-3 *ATP Control of Protein Function: Loading a Sliding Clamp*, 282

DNA SYNTHESIS AT THE REPLICATION FORK, 283

Interactions between Replication Fork Proteins Form the *E. coli* Replisome, 286

INITIATION OF DNA REPLICATION, 288

Specific Genomic DNA Sequences Direct the Initiation of DNA Replication, 288

The Replicon Model of Replication Initiation, 288

Replicator Sequences Include Initiator-Binding Sites and Easily Unwound DNA, 289

KEY EXPERIMENTS BOX 9-4 *The Identification of Origins of Replication and Replicators*, 290

BINDING AND UNWINDING: ORIGIN SELECTION AND ACTIVATION BY THE INITIATOR PROTEIN, 293

Protein-Protein and Protein-DNA Interactions Direct the Initiation Process, 293

ADVANCED CONCEPTS BOX 9-5 *E. coli DNA Replication Is Regulated by DnaA-ATP Levels and SeqA*, 294

Eukaryotic Chromosomes Are Replicated Exactly Once per Cell Cycle, 297

Helicase Loading Is the First Step in the Initiation of Replication in Eukaryotes, 298

Helicase Loading and Activation Are Regulated to Allow Only a Single Round of Replication during Each Cell Cycle, 300