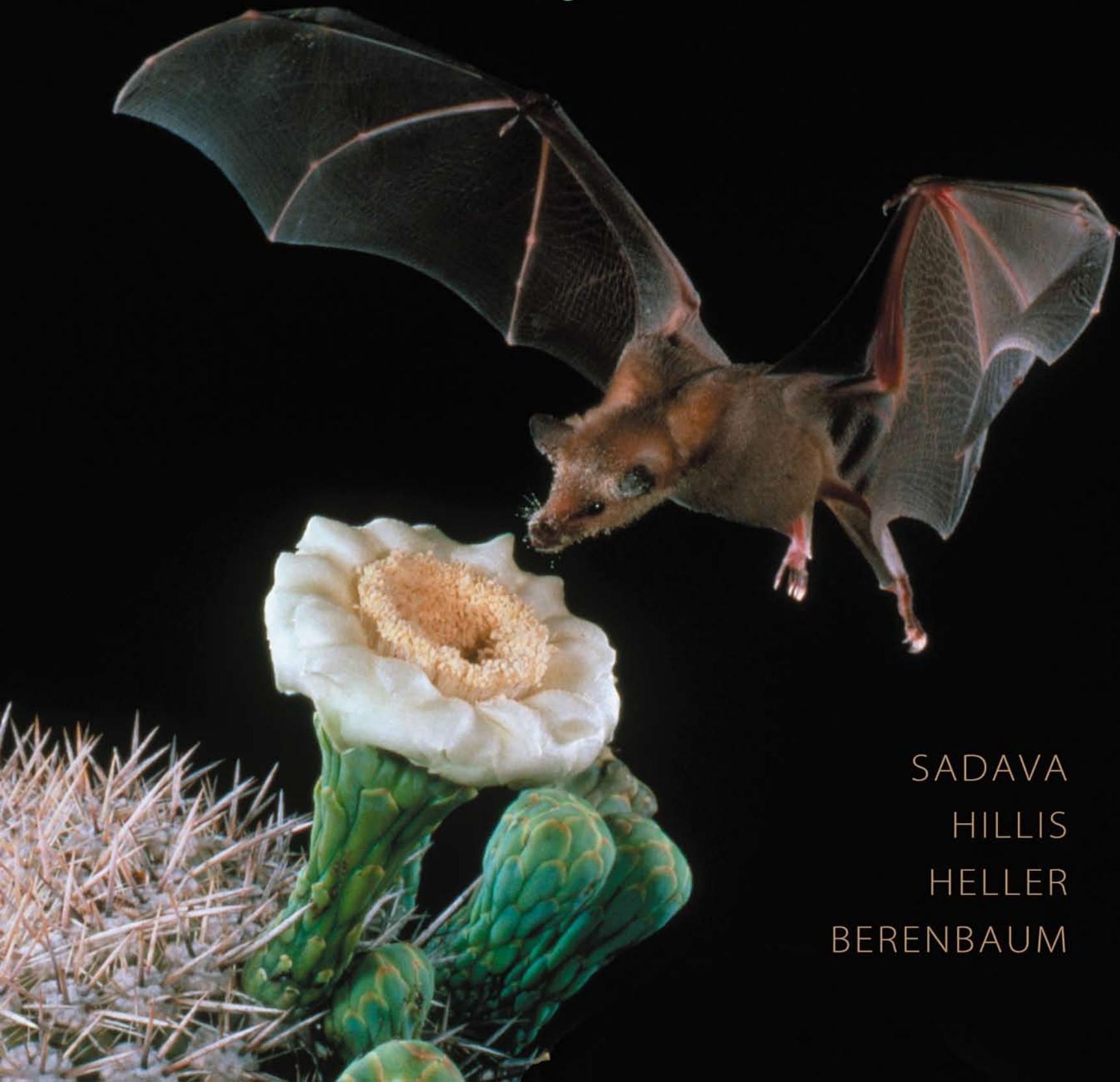


Life

The Science
of Biology
NINTH EDITION



SADAVA
HILLIS
HELLER
BERENBAUM

Life

The Science of Biology

NINTH EDITION



Sinauer Associates, Inc.



W. H. Freeman and Company

NINTH
EDITION

Life

*The Science
of Biology*

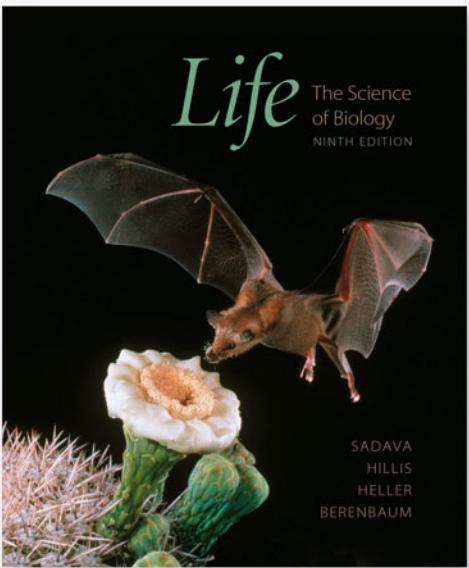


DAVID
SADAVA
The Claremont Colleges
Claremont, California

DAVID M.
HILLIS
University of Texas
Austin, Texas

H. CRAIG
HELLER
Stanford University
Stanford, California

MAY R.
BERENBAUM
University of Illinois
Urbana-Champaign, Illinois



About the Cover

The cover of *Life* captures many themes that echo throughout the book. The photograph shows a lesser long-nosed bat pollinating a saguaro cactus. This cactus has evolved large flowers that produce copious quantities of nectar. The nectar attracts many species that pollinate the cactus, including bats. The ability of bats to hover as they feed on the nectar of the cactus is an excellent example of adaptation of body form and physiology. These themes of adaptation, evolution, nutrition, reproduction, species interactions, and integrated form and function are ideas that are repeated throughout the chapters of *Life*. Photograph copyright © Dr. Merlin D. Tuttle/Photo Researchers, Inc.

The Frontispiece

Blue wildebeest and Burchell's zebra migrate together through Serengeti National Park, Tanzania. Copyright © Art Wolfe, www.artwolfe.com.

LIFE: The Science of Biology, Ninth Edition

Copyright © 2011 by Sinauer Associates, Inc. All rights reserved.

This book may not be reproduced in whole or in part without permission.

Address editorial correspondence to:

Sinauer Associates Inc., 23 Plumtree Road, Sunderland, MA 01375 U.S.A.

www.sinauer.com

publish@sinauer.com

Address orders to:

MPS / W. H. Freeman & Co., Order Dept., 16365 James Madison Highway,
U.S. Route 15, Gordonsville, VA 22942 U.S.A.

Examination copy information: 1-800-446-8923

Orders: 1-888-330-8477



Library of Congress Cataloging-in-Publication Data

Life, the science of biology / David Sadava .. [et al.]. — 9th ed.

p. cm.

Includes index.

ISBN 978-1-4292-1962-4 (hardcover) — 978-1-4292-4645-3 (pbk. : v. 1) —

ISBN 978-1-4292-4644-6 (pbk. : v. 2) — ISBN 978-1-4292-4647-7 (pbk. : v. 3)

1. Biology. I. Sadava, David E.

QH308.2.L565 2011

570—dc22

2009036693

Printed in U.S.A.

First Printing October 2009

The Courier Companies, Inc.

*To Bill Purves and Gordon Orians,
extraordinary colleagues, biologists, and teachers,
and the original authors of LIFE*

The Authors



CRAIG HELLER

DAVID HILLIS

MAY BERENBAUM

DAVID SADAVA

DAVID SADAVA is the Pritzker Family Foundation Professor of Biology, Emeritus, at the Keck Science Center of Claremont McKenna, Pitzer, and Scripps, three of The Claremont Colleges. In addition, he is Adjunct Professor of Cancer Cell Biology at the City of Hope Medical Center. Twice winner of the Huntoon Award for superior teaching, Dr. Sadava taught courses on introductory biology, biotechnology, biochemistry, cell biology, molecular biology, plant biology, and cancer biology. In addition to *Life: The Science of Biology*, he is the author or coauthor of books on cell biology and on plants, genes, and crop biotechnology. His research has resulted in many papers coauthored with his students, on topics ranging from plant biochemistry to pharmacology of narcotic analgesics to human genetic diseases. For the past 15 years, he has investigated multi-drug resistance in human small-cell lung carcinoma cells with a view to understanding and overcoming this clinical challenge. At the City of Hope, his current work focuses on new anti-cancer agents from plants.

DAVID HILLIS is the Alfred W. Roark Centennial Professor in Integrative Biology and the Director of the Center for Computational Biology and Bioinformatics at the University of Texas at Austin, where he also has directed the School of Biological Sciences. Dr. Hillis has taught courses in introductory biology, genetics, evolution, systematics, and biodiversity. He has been elected into the membership of the National Academy of Sciences and the American Academy of Arts and Sciences, awarded a John D. and Catherine T. MacArthur Fellowship, and has served as President of the Society for the Study of Evolution and of the Society of Systematic Biologists. His research interests span much of evolutionary biology, including experimental studies of evolving viruses, empirical studies of natural molecular evolution, applications of phylogenetics, analyses of biodiversity, and evolutionary modeling. He is particularly interested in teaching and research about the practical applications of evolutionary biology.

CRAIG HELLER is the Lorry I. Lokey /BusinessWire Professor in Biological Sciences and Human Biology at Stanford University. He earned his Ph.D. from the Department of Biology at Yale University in 1970. Dr. Heller has taught in the core biology courses at Stanford since 1972 and served as Director of the Program in Human Biology, Chairman of the Biological Sciences Department, and Associate Dean of Research. Dr. Heller is a fellow of the American Association for the Advancement of Science and a recipient of the Walter J. Gores Award for excellence in teaching. His research is on the neurobiology of sleep and circadian rhythms, mammalian hibernation, the regulation of body temperature, the physiology of human performance, and the neurobiology of learning. Dr. Heller has done research on a huge variety of animals and physiological problems ranging from sleeping kangaroo rats, diving seals, hibernating bears, photoperiodic hamsters, and exercising athletes. Some of his recent work on the effects of temperature on human performance is featured in the opener to Chapter 40, "Physiology, Homeostasis, and Temperature Regulation."

MAY BERENBAUM is the Swanlund Professor and Head of the Department of Entomology at the University of Illinois at Urbana-Champaign. She has taught courses in introductory animal biology, entomology, insect ecology, and chemical ecology, and has received awards at the regional and national level for distinguished teaching from the Entomological Society of America. A fellow of the National Academy of Sciences, the American Academy of Arts and Sciences, and the American Philosophical Society, she served as President of the American Institute for Biological Sciences in 2009. Her research addresses insect–plant coevolution, from molecular mechanisms of detoxification to impacts of herbivory on community structure. Concerned with the practical application of ecological and evolutionary principles, she has examined impacts of genetic engineering, global climate change, and invasive species on natural and agricultural ecosystems. Devoted to fostering science literacy, she has published numerous articles and five books on insects for the general public.

Contents in Brief

PART ONE ■ THE SCIENCE OF LIFE AND ITS CHEMICAL BASIS

- 1 Studying Life 2
- 2 Small Molecules and the Chemistry of Life 20
- 3 Proteins, Carbohydrates, and Lipids 38
- 4 Nucleic Acids and the Origin of Life 60

PART TWO ■ CELLS

- 5 Cells: The Working Units of Life 76
- 6 Cell Membranes 105
- 7 Cell Signaling and Communication 128

PART THREE ■ CELLS AND ENERGY

- 8 Energy, Enzymes, and Metabolism 148
- 9 Pathways that Harvest Chemical Energy 168
- 10 Photosynthesis: Energy from Sunlight 189

PART FOUR ■ GENES AND HEREDITY

- 11 The Cell Cycle and Cell Division 209
- 12 Inheritance, Genes, and Chromosomes 236
- 13 DNA and Its Role in Heredity 266
- 14 From DNA to Protein: Gene Expression 290
- 15 Gene Mutation and Molecular Medicine 316
- 16 Regulation of Gene Expression 342

PART FIVE ■ GENOMES

- 17 Genomes 365
- 18 Recombinant DNA and Biotechnology 386
- 19 Differential Gene Expression in Development 405
- 20 Development and Evolutionary Change 426

PART SIX ■ THE PATTERNS AND PROCESSES OF EVOLUTION

- 21 Evidence and Mechanisms of Evolution 440
- 22 Reconstructing and Using Phylogenies 464
- 23 Species and Their Formation 481
- 24 Evolution of Genes and Genomes 498
- 25 The History of Life on Earth 518

PART SEVEN ■ THE EVOLUTION OF DIVERSITY

- 26 Bacteria and Archaea: The Prokaryotic Domains 536
- 27 The Origin and Diversification of Eukaryotes 560
- 28 Plants without Seeds: From Water to Land 588

- 29 The Evolution of Seed Plants 607
- 30 Fungi: Recyclers, Pathogens, Parasites, and Plant Partners 626
- 31 Animal Origins and the Evolution of Body Plans 645
- 32 Protostome Animals 666
- 33 Deuterostome Animals 692

PART EIGHT ■ FLOWERING PLANTS: FORM AND FUNCTION

- 34 The Plant Body 719
- 35 Transport in Plants 739
- 36 Plant Nutrition 755
- 37 Regulation of Plant Growth 771
- 38 Reproduction in Flowering Plants 794
- 39 Plant Responses to Environmental Challenges 814

PART NINE ■ ANIMALS: FORM AND FUNCTION

- 40 Physiology, Homeostasis, and Temperature Regulation 832
- 41 Animal Hormones 851
- 42 Immunology: Animal Defense Systems 873
- 43 Animal Reproduction 899
- 44 Animal Development 922
- 45 Neurons and Nervous Systems 943
- 46 Sensory Systems 964
- 47 The Mammalian Nervous System: Structure and Higher Function 985
- 48 Musculoskeletal Systems 1006
- 49 Gas Exchange in Animals 1025
- 50 Circulatory Systems 1045
- 51 Nutrition, Digestion, and Absorption 1067
- 52 Salt and Water Balance and Nitrogen Excretion 1091
- 53 Animal Behavior 1113

PART TEN ■ ECOLOGY

- 54 Ecology and the Distribution of Life 1140
- 55 Population Ecology 1167
- 56 Species Interaction and Coevolution 1185
- 57 Community Ecology 1203
- 58 Ecosystems and Global Ecology 1221
- 59 Conservation Biology 1242

Investigating Life/Tools for Investigating Life

INVESTIGATING LIFE

- 1.13** Controlled Experiments Manipulate a Variable 15
1.14 Comparative Experiments Look for Differences among Groups 16
3.9 Primary Structure Specifies Tertiary Structure 47
4.7 Disproving the Spontaneous Generation of Life 66
4.9 Miller and Urey Synthesized Prebiotic Molecules in an Experimental Atmosphere 68
5.23 The Role of Microfilaments in Cell Movement—Showing Cause and Effect in Biology 99
6.5 Rapid Diffusion of Membrane Proteins 110
6.13 Aquaporin Increases Membrane Permeability to Water 119
7.13 The Discovery of a Second Messenger 138
9.10 Two Experiments Demonstrate the Chemiosmotic Mechanism 180
10.2 The Source of the Oxygen Produced by Photosynthesis 191
10.13 Tracing the Pathway of CO₂ 199
11.4 Regulation of the Cell Cycle 213
12.3 Mendel’s Monohybrid Experiments 241
12.6 Homozygous or Heterozygous? 243
12.18 Some Alleles Do Not Assort Independently 254
13.1 Genetic Transformation 268
13.2 Genetic Transformation by DNA 269
13.4 The Hershey–Chase Experiment 271
13.5 Transfection in Eukaryotic Cells 272
13.11 The Meselson–Stahl Experiment 277
- 14.1** One Gene, One Enzyme 293
14.5 Deciphering the Genetic Code 298
14.9 Demonstrating the Existence of Introns 302
14.20 Testing the Signal 311
15.23 Gene Therapy: The Ex Vivo Approach 338
17.8 Using Transposon Mutagenesis to Determine the Minimal Genome 375
18.1 Recombinant DNA 388
19.2 Developmental Potential in Early Frog Embryos 408
19.3 Cloning a Plant 409
19.8 Asymmetry in the Early Sea Urchin Embryo 414
20.7 Changing the Form of an Appendage 431
20.11 Temperature-Dependent Sex Determination is Associated with Sex-Specific Fitness Differences 434
21.16 Sexual Selection in Male Widowbirds 454
21.17 Do Bright Bills Signal Good Health? 455
21.19 A Heterozygote Mating Advantage 458
22.7 The Accuracy of Phylogenetic Analysis 472
23.13 Flower Color Reinforces a Reproductive Barrier in *Phlox* 492
24.4 Evolution in a Heterogeneous Environment 504
25.6 Rising Oxygen Levels and Body Size in Insects 524
26.21 What Is the Highest Temperature Compatible with Life? 551
27.6 The Role of Vacuoles in Ciliate Digestion 568
28.9 CO₂ Levels and the Evolution of Megaphylls 598
29.14 Stigma Behavior in *Mimulus* Flowers 618
- 35.10** Measuring Potassium Ion Concentration in Guard Cells 749
36.3 Is Nickel an Essential Element for Plant Growth? 759
37.8 The Darwins’ Phototropism Experiment 780
37.9 Went’s Experiment 780
37.20 Sensitivity of Seeds to Red and Far-Red Light 789
38.13 Night Length and Flowering 805
38.14 Interrupting the Night 806
38.15 The Flowering Signal Moves from Leaf to Bud 806
39.7 Nicotine Is a Defense against Herbivores 821
39.17 A Molecular Response to Drought Stress 827
40.19 The Hypothalamus Regulates Body Temperature 847
41.3 A Diffusible Substance Triggers Molting 855
42.5 The Discovery of Specific Immunity 881
43.16 Prolactin and the Male Refractory Period 917
44.10 The Dorsal Lip Induces Embryonic Organization 931
44.12 Differentiation Can Be Due to Inhibition of Transcription Factors 933
45.17 Repeated Stimulation Can Cause Long-Term Potentiation 961
46.14 A Rod Cell Responds to Light 977
47.11 What Does the Eye Tell the Brain? 996
47.12 Cells in the Visual Cortex Respond to Specific Patterns of Light 997
48.8 Neurotransmitters Alter the Membrane Potential of Smooth Muscle Cells 1013
49.17 Sensitivity of the Respiratory Control System Changes with Exercise 1041

- | | | |
|---|--|--|
| 50.5 The Autonomic Nervous System Controls Heart Rate 1054 | 53.10 Flycatchers Use Neighbors' Success to Assess Habitat Quality 1123 | 56.14 Are Ants and Acacias Mutualists? 1197 |
| 51.19 A Single-Gene Mutation Leads to Obesity in Mice 1087 | 53.11 The Costs of Defending a Territory 1124 | 57.10 Testing the Theory of Island Biogeography 1213 |
| 52.12 An Ammonium Transporter in the Renal Tubules? 1105 | 53.13 Bluegill Sunfish Are Energy Maximizers 1126 | 58.18 Effects of Atmospheric CO ₂ Concentrations on Nitrogen Fixation 1237 |
| 52.17 ADH Induces Insertion of Aquaporins into Plasma Membranes 1109 | 53.18 A Time-Compensated Solar Compass 1131 | 59.14 Species Richness Enhances Wetland Restoration 1253 |
| | 55.11 Corridors Can Affect Metapopulations 1179 | |

TOOLS FOR INVESTIGATING LIFE

- | | | |
|---|--|--|
| 5.3 Looking at Cells 80–81 | 17.1 Sequencing Genomes Involves Fragment Overlaps 367 | 35.8 Measuring the Pressure of Xylem Sap with a Pressure Chamber 747 |
| 5.6 Cell Fractionation 85 | 17.2 Sequencing DNA 368 | 36.2 Growing Plants Hydroponically 758 |
| 6.4 Membrane Proteins Revealed by the Freeze-Fracture Technique 109 | 17.3 High-Throughput Sequencing 369 | 37.3 A Genetic Screen 776 |
| 12.2 A Controlled Cross between Two Plants 239 | 18.4 Marking Recombinant DNA by Inactivating a Gene 391 | 41.17 An Immunoassay Allows Measurement of Small Concentrations 869 |
| 14.8 Nucleic Acid Hybridization 301 | 18.6 Constructing Libraries 392 | 42.10 Creating Hybridomas for the Production of Monoclonal Antibodies 887 |
| 15.8 Separating Fragments of DNA by Gel Electrophoresis 325 | 18.7 Making a Knockout Mouse 394 | 45.5 Measuring the Membrane Potential 949 |
| 15.15 RFLP Markers 331 | 18.9 DNA on a Chip 396 | 45.7 Using the Nernst Equation 950 |
| 15.16 Analyzing DNA Fragments by DNA Gel Blotting 332 | 19.4 Cloning a Mammal 410 | 45.8 Patch Clamping 951 |
| 15.20 DNA Testing by Allele-Specific Cleavage 336 | 21.6 Calculating Allele Frequencies 446 | 49.9 Measuring Lung Ventilation 1033 |
| 15.21 DNA Testing by Allele-Specific Oligonucleotide Hybridization 336 | 24.1 Amino Acid Sequence Alignment 500 | |
| | 26.4 The Microchemostat 540 | |

Preface

Biology is a dynamic, exciting, and important subject. It is dynamic because it is constantly changing, with new discoveries about the living world being made every day. (Although it is impossible to pinpoint an exact number, approximately 1 million new research articles in biology are published each year.) The subject is exciting because life in all of its forms has always fascinated people. As active scientists who have spent our careers teaching and doing research in a wide variety of fields, we know this first hand.

Biology has always been important in peoples' daily lives, if only through the effects of achievements in medicine and agriculture. Today more than ever the science of biology is at the forefront of human concerns as we face challenges raised both by recent advances in genome science and by the rapidly changing environment.

Life's new edition brings a fresh approach to the study of biology while retaining the features that have made the book successful in the past. A new coauthor, the distinguished entomologist May R. Berenbaum (University of Illinois at Urbana-Champaign) has joined our team, and the role of evolutionary biologist David Hillis (University of Texas at Austin) is greatly expanded in this edition. The authors hail from large, medium-sized, and small institutions. Our multiple perspectives and areas of expertise, as well as input from many colleagues and students who used previous editions, have informed our approach to this new edition.

Enduring Features

We remain committed to blending the presentation of core ideas with an emphasis on introducing students to the *process of scientific inquiry*. Having pioneered the idea of depicting seminal experiments in specially designed figures, we continue to develop this here, with 79 **INVESTIGATING LIFE** figures. Each of these figures sets the experiment in perspective and relates it to the accompanying text. As in previous editions, these figures employ a structure: Hypothesis, Method, Results, and Conclusion. They often include questions for further research that ask students to conceive an experiment that would explore a related question. Each *Investigating Life* figure has a reference to BioPortal (yourBioPortal.com), where citations to the original work as well as additional discussion and references to follow-up research can be found.

A related feature is the **TOOLS FOR INVESTIGATING LIFE** figures, which depict laboratory and field methods used in biology. These, too, have been expanded to provide more useful context for their importance.

Over a decade ago—in *Life's* Fifth Edition—the authors and publishers pioneered the much-praised use of **BALLOON CAPTIONS** in our figures. We recognized then, and it is even truer today, that many students are visual learners. The balloon cap-

tions bring explanations of intricate, complex processes directly into the illustration, allowing students to integrate information without repeatedly going back and forth between the figure, its legend, and the text.

Life is the only introductory textbook for biology majors to begin each chapter with a story. These **OPENING STORIES** provide historical, medical, or social context and are intended to intrigue students while helping them see how the chapter's biological subject relates to the world around them. In the new edition, all of the opening stories (some 70 percent of which are new) are revisited in the body of the chapter to drive home their relevance.

We continue to refine our well-received *chapter organization*. The chapter-opening story ends with a brief **IN THIS CHAPTER** preview of the major subjects to follow. A **CHAPTER OUTLINE** asks questions to emphasize scientific inquiry, each of which is answered in a major section of the chapter. A **RECAP** at the end of each section asks the student to pause and answer questions to review and test their mastery of the previous material. The end-of-chapter summary continues this inquiry framework and highlights key figures, bolded terms, and activities and animated tutorials available in BioPortal.

New Features

Probably the most important new feature of this edition is *new authorship*. Like the biological world, the authorship team of *Life* continues to evolve. While two of us (Craig Heller and David Sadava) continue as coauthors, David Hillis has a greatly expanded role, with full responsibility for the units on evolution and diversity. New coauthor May Berenbaum has rewritten the chapters on ecology. The perspectives of these two acclaimed experts have invigorated the entire book (as well as their coauthors).

Even with the enduring features (see above), this edition has a different look and feel from its predecessor. A fresh *new design* is more open and, we hope, more accessible to students. The extensively *revised art program* has a contemporary style and color palette. The information flow of the figures is easier to follow, with numbered balloons as a guide for students. There are new conceptual figures, including a striking visual timeline for the evolution of life on Earth (Figure 25.12) and a single overview figure that summarizes the information in the genome (Figure 17.4).

In response to instructors who asked for more real-world data, we have incorporated a feature introduced online in the Eighth Edition, **WORKING WITH DATA**. There are now 36 of these exercises, most of which relate to an *Investigating Life* figure. Each is referenced at the end of the relevant chapter and is available online via BioPortal (yourBioPortal.com). In these exercises, we describe in detail the context and approach of the

research paper that forms the basis of the figure. We then ask the student to examine the data, to make calculations, and to draw conclusions.

We are proud that this edition is a *greener Life*, with the goal of reducing our environmental impact. This is the first introductory biology text to be printed on paper earning the Forest Stewardship Council label, the “gold standard” in green paper products, and it is manufactured from wood harvested from sustainable forests. And, of course, we also offer *Life* as an eBook.

The Ten Parts

We have reorganized the book into ten parts. **Part One, The Science of Life and Its Chemical Basis**, sets the stage for the book: the opening chapter focuses on biology as an exciting science. We begin with a startling observation: the recent, dramatic decline of amphibian species throughout the world. We then show how biologists have formed hypotheses for the causes of this environmental problem and are testing them by carefully designed experiments, with a view not only to understanding the decline, but reversing it. This leads to an outline of the basic principles of biology that are the foundation for the rest of the book: the unity of life at the cellular level and how evolution unites the living world. This is followed by chapters on the basic chemical building blocks that underlie life. We have added a new chapter on nucleic acids and the origin of life, introducing the concepts of genes and gene expression early and expanding our coverage of the major ideas on how life began and evolved at its earliest stages.

In **Part Two, Cells**, we describe the view of life as seen through cells, its structural units. In response to comments by users of our previous edition, we have moved the chapter on cell signaling and communication from the genetics section to this part of the book, with a change in emphasis from genes to cells. There is an updated discussion of ideas on the origin of cells and organelles, as well as expanded treatment of water transport across membranes.

Part Three, Cells and Energy, presents an integrated view of biochemistry. For this edition, we have worked to clarify such challenging concepts as energy transfer, allosteric enzymes, and biochemical pathways. There is extensive revision of the discussions of alternate pathways of photosynthetic carbon fixation, as well as a greater emphasis on applications throughout these chapters.

Part Four, Genes and Heredity, is extensively revised and reorganized to improve clarity, link related concepts, and provide updates from recent research results. Separate chapters on prokaryotic genetics and molecular medicine have been removed and their material woven into relevant chapters. For example, our chapter on cell reproduction now includes a discussion of how the basic mechanisms of cell division are altered in cancer cells. The chapter on transmission genetics now includes coverage of this phenomenon in prokaryotes. New chapters on gene expression and gene regulation compare prokaryotic and eukaryotic mechanisms and include a discussion of

epigenetics. A new chapter on mutation describes updated applications of medical genetics.

In **Part Five, Genomes**, we reinforce the concepts of the previous part, beginning with a new chapter on genomes—how they are analyzed and what they tell us about the biology of prokaryotes and eukaryotes, including humans. This leads to a chapter describing how our knowledge of molecular biology and genetics underpins biotechnology (the application of this knowledge to practical problems). We discuss some of the latest uses of biotechnology, including environmental cleanup. Part Five finishes with two chapters on development that explore the themes of molecular biology and evolution, linking these two parts of the book.

Part Six, The Patterns and Processes of Evolution, emphasizes the importance of evolutionary biology as a basis for comparing and understanding all aspects of biology. These chapters have been extensively reorganized and revised, as well as updated with the latest thinking of biologists in this rapidly changing field. This part now begins with the evidence and mechanisms of evolution, moves into a discussion of phylogenetic trees, then covers speciation and molecular evolution, and concludes with the evolutionary history of life on Earth. An integrated timeline of evolutionary history shows the timing of major events of biological evolution, the movements of the continents, floral and faunal reconstructions of major time periods, and depicts some of the fossils that form the basis of the reconstructions.

In **Part Seven, The Evolution of Diversity**, we describe the latest views on biodiversity and evolutionary relationships. Each chapter has been revised to make it easier for the reader to appreciate the major changes that have evolved within the various groups of organisms. We emphasize understanding the big picture of organismal diversity, as opposed to memorizing a taxonomic hierarchy and names (although these are certainly important). Throughout the book, the tree of life is emphasized as a way of understanding and organizing biological information. A *Tree of Life Appendix* allows students to place any group of organisms mentioned in the text of our book into the context of the rest of life. The web-based version of this appendix provides links to photos, keys, species lists, distribution maps, and other information to help students explore biodiversity of specific groups in greater detail.

After modest revisions in the past two editions, **Part Eight, Flowering Plants: Form and Function**, has been extensively reorganized and updated with the help of Sue Wessler, to include both classical and more recent approaches to plant physiology. Our emphasis is not only on the basic findings that led to the elucidation of mechanisms for plant growth and reproduction, but also on the use of genetics of model organisms. There is expanded coverage of the cell signaling events that regulate gene expression in plants, integrating concepts introduced earlier in the book. New material on how plants respond to their environment is included, along with links to both the book’s earlier descriptions of plant diversity and later discussions of ecology.

Part Nine, Animals: Form and Function, continues to provide a solid foundation in physiology through comprehensive coverage of basic principles of function of each organ system and then emphasis on mechanisms of control and integration. An important reorganization has been moving the chapter on immunology from earlier in the book, where its emphasis was on molecular genetics, to this part, where it is more closely allied to the information systems of the body. In addition, we have added a number of new experiments and made considerable effort to clarify the sometimes complex phenomena shown in the illustrations.

Part Ten, Ecology, has been significantly revised by our new coauthor, May Berenbaum. A new chapter of biological interactions has been added (a topic formerly covered in the community ecology chapter). Full of interesting anecdotes and discussions of field studies not previously described in biology texts, this new ecology unit offers practical insights into how ecologists acquire, interpret, and apply real data. This brings the book full circle, drawing upon and reinforcing prior topics of energy, evolution, phylogenetics, Earth history, and animal and plant physiology.

Exceptional Value Formats

We again provide *Life* both as the full book and as a cluster of paperbacks. Thus, instructors who want to use less than the whole book can choose from these split volumes, each with the book's front matter, appendices, glossary, and index.

Volume I, The Cell and Heredity, includes: Part One, The Science of Life and Its Chemical Basis (Chapters 1–4); Part Two, Cells (Chapters 5–7); Part Three, Cells and Energy (Chapters 8–10); Part Four, Genes and Heredity (Chapters 11–16); and Part Five, Genomes (Chapters 17–20).

Volume II, Evolution, Diversity, and Ecology, includes: Chapter 1, Studying Life; Part Six, The Patterns and Processes of Evolution (Chapters 21–25); Part Seven, The Evolution of Diversity (Chapters 26–33); and Part Ten, Ecology (Chapters 54–59).

Volume III, Plants and Animals, includes: Chapter 1, Studying Life; Part Eight, Flowering Plants: Form and Function (Chapters 34–39); and Part Nine, Animals: Form and Function (Chapters 40–53).

Responding to student concerns, we offer two options of the entire book at a *significantly reduced cost*. After it was so well received in the previous edition, we again provide *Life* as a *loose-leaf version*. This shrink-wrapped, unbound, 3-hole punched version fits into a 3-ring binder. Students take only what they need to class and can easily integrate any instructor handouts or other resources.

Life was the first comprehensive biology text to offer the entire book as a truly robust *eBook*. For this edition, we continue to offer a flexible, interactive ebook that gives students a new way to read the text and learn the material. The ebook integrates the student media resources (animations, quizzes, activities, etc.) and offers instructors a powerful way to customize the textbook with their own text, images, Web links, documents, and more.

Media and Supplements for the Ninth Edition

The wide range of media and supplements that accompany *Life*, Ninth Edition have all been created with the dual goal of helping students learn the material presented in the textbook more efficiently and helping instructors teach their courses more effectively. Students in majors introductory biology are faced with learning a tremendous number of new concepts, facts, and terms, and the more different ways they can study this material, the more efficiently they can master it.

All of the *Life* media and supplemental resources have been developed specifically for this textbook. This provides strong consistency between text and media, which in turn helps students learn more efficiently. For example, the animated tutorials and activities found in BioPortal were built using textbook art, so that the manner in which structures are illustrated, the colors used to identify objects, and the terms and abbreviations used are all consistent.

For the Ninth Edition, a new set of Interactive Tutorials gives students a new way to explore many key topics across the textbook. These new modules allow the student to learn by doing, including solving problem scenarios, working with experimental techniques, and exploring model systems. All new copies of the Ninth Edition include access to the robust new version of BioPortal, which brings together all of *Life*'s student and instructor resources, powerful assessment tools, and new integration with Prep-U adaptive quizzing.

The rich collection of visual resources in the Instructor's Media Library provides instructors with a wide range of options for enhancing lectures, course websites, and assignments. Highlights include: layered art PowerPoint® presentations that break down complex figures into detailed, step-by-step presentations; a collection of approximately 200 video segments that can help capture the attention and imagination of students; and PowerPoint slides of textbook art with editable labels and leaders that allow easy customization of the figures.

For a detailed description of all the media and supplements available for the Ninth Edition, please turn to "Life's Media and Supplements," on page xvii.

Many People to Thank

"If I have seen farther, it is by standing on the shoulders of giants." The great scientist Isaac Newton wrote these words over 330 years ago and, while we certainly don't put ourselves in his lofty place in science, the words apply to us as coauthors of this text. This is the first edition that does not bear the names of Bill Purves and Gordon Orians. As they enjoy their "retirements," we are humbled by their examples as biologists, educators, and writers.

One of the wisest pieces of advice ever given to a textbook author is to "be passionate about your subject, but don't put your ego on the page." Considering all the people who looked over our shoulders throughout the process of creating this book, this advice could not be more apt. We are indebted to many people who gave invaluable help to make this book what it is. First and foremost are our colleagues, biologists from over 100 institutions. Some were users of the previous edition, who suggested many improvements. Others reviewed our chapter drafts in detail, including advice on how to improve the illustrations. Still others acted as accuracy reviewers when the book was almost completed. All of these biologists are listed in the Reviewer credits.

Of special note is Sue Wessler, a distinguished plant biologist and textbook author from the University of Georgia. Sue looked critically at Part Eight, Flowering Plants: Form and Function, wrote three of the chapters (34–36), and was important in the revision of the other three (37–39). The new approach to plant biology in this edition owes a lot to her.

The pace of change in biology and the complexities of preparing a book as broad as this one necessitated having two developmental editors. James Funston coordinated Parts 1–5, and Carol Pritchard-Martinez coordinated Parts 6–10. We benefitted from the wide experience, knowledge, and wisdom of both of them. As the chapter drafts progressed, we were fortunate to have experienced biologist Laura Green lending her critical eye as in-house editor. Elizabeth Morales, our artist, was on her third edition with us. As we have noted, she extensively revised almost all of the prior art and translated our crude sketches into beautiful new art. We hope you agree that our art program re-

mains superbly clear and elegant. Our copy editors, Norma Roche, Liz Pierson, and Jane Murfett, went far beyond what such people usually do. Their knowledge and encyclopedic recall of our book's chapters made our prose sharper and more accurate. Diane Kelly, Susan McGlew, and Shannon Howard effectively coordinated the hundreds of reviews that we described above. David McIntyre was a terrific photo editor, finding over 550 new photographs, including many new ones of his own, that enrich the book's content and visual statement. Jefferson Johnson is responsible for the design elements that make this edition of *Life* not just clear and easy to learn from, but beautiful as well. Christopher Small headed the production department—Joanne Delphia, Joan Gemme, Janice Holabird, and Jefferson Johnson—who contributed in innumerable ways to bringing *Life* to its final form. Jason Dirks once again coordinated the creation of our array of media and supplements, including our superb new Web resources. Carol Wigg, for the ninth time in nine editions, oversaw the editorial process; her influence pervades the entire book.

W. H. Freeman continues to bring *Life* to a wider audience. Associate Director of Marketing Debbie Clare, the Regional Specialists, Regional Managers, and experienced sales force are effective ambassadors and skillful transmitters of the features and unique strengths of our book. We depend on their expertise and energy to keep us in touch with how *Life* is perceived by its users. And thanks also to the Freeman media group for eBook and BioPortal production.

Finally, we are indebted to Andy Sinauer. Like ours, his name is on the cover of the book, and he truly cares deeply about what goes into it. Combining decades of professionalism, high standards, and kindness to all who work with him, he is truly our mentor and friend.

DAVID SADAVA

DAVID HILLIS

CRAIG HELLER

MAY BERENBAUM

Reviewers for the Ninth Edition

Between-Edition Reviewers

David D. Ackerly, University of California, Berkeley
Amy Bickham Baird, University of Leiden
Jeremy Brown, University of California, Berkeley
John M. Burke, University of Georgia
Ruth E. Buskirk, University of Texas, Austin
Richard E. Duhrkopf, Baylor University
Casey W. Dunn, Brown University
Erika J. Edwards, Brown University
Kevin Folta, University of Florida
Lynda J. Goff, University of California, Santa Cruz
Tracy A. Heath, University of Kansas
Shannon Hettke, University of Texas, Austin
Richard H. Heineman, University of Texas, Austin
Albert Herrera, University of Southern California
David S. Hibbett, Clark University
Norman A. Johnson, University of Massachusetts
Walter S. Judd, University of Florida
Laura A. Katz, Smith College
Emily Moriarty Lemmon, Florida State University
Sheila McCormick, University of California, Berkeley
Robert McCurdy, Independence Creek Nature Preserve
Jacalyn Newman, University of Pittsburgh
Juliet F. Noor, Duke University
Theresa O'Halloran, University of Texas, Austin
K. Sata Sathasivan, University of Texas, Austin
H. Bradley Shaffer, University of California, Davis
Rebecca Symula, Yale University
Christopher D. Todd, University of Saskatchewan
Elizabeth Willott, University of Arizona
Kenneth Wilson, University of Saskatchewan

Manuscript Reviewers

Tamarah Adair, Baylor University
William Adams, University of Colorado, Boulder
Gladys Alexandre, University of Tennessee, Knoxville
Shivanthi Anandan, Drexel University
Brian Bagatto, University of Akron
Lisa Baird, University of San Diego
Stewart H. Berlocher, University of Illinois, Urbana-Champaign
William Bischoff, University of Toledo
Meredith M. Blackwell, Louisiana State University
David Bos, Purdue University
Jonathan Bossenbroek, University of Toledo
Nicole Bournias-Vardiabasis, California State University, San Bernardino
Nancy Boury, Iowa State University
Sunny K. Boyd, University of Notre Dame
Judith L. Bronstein, University of Arizona
W. Randy Brooks, Florida Atlantic University
James J. Bull, University of Texas, Austin
Darlene Campbell, Cornell University
Domenic Castignetti, Loyola University, Chicago
David T. Champlin, University of Southern Maine
Shu-Mei Chang, University of Georgia
Samantha K. Chapman, Villanova University
Patricia Christie, MIT
Wes Colgan, Pikes Peak Community College
John Cooper, Washington University
Ronald Cooper, University of California, Los Angeles
Elizabeth Cowles, Eastern Connecticut State University
Jerry Coyne, University of Chicago
William Crampton, University of Central Florida
Michael Dalbey, University of California, Santa Cruz
Anne Danielson-Francois, University of Michigan, Dearborn

Grayson S. Davis, Valparaiso University
Kevin Dixon, Florida State University
Zaldy Doyungan, Texas A&M University, Corpus Christi
Ernest F. Dubrul, University of Toledo
Roland Dute, Auburn University
Scott Edwards, Harvard University
William Eldred, Boston University
David Eldridge, Baylor University
Joanne Ellzey, University of Texas, El Paso
Susan H. Erster, State University of New York, Stony Brook
Brent E. Ewers, University of Wyoming
Kevin Folta, University of Florida
Brandon Foster, Wake Technical Community College
Richard B. Gardiner, University of Western Ontario
Douglas Gayou, University of Missouri, Columbia
John R. Geiser, Western Michigan University
Arundhati Ghosh, University of Pittsburgh
Alice Gibb, Northern Arizona University
Scott Gilbert, Swarthmore College
Matthew R. Gilg, University of North Florida
Elizabeth Godrick, Boston University
Lynda J. Goff, University of California, Santa Cruz
Elizabeth Blinstrup Good, University of Illinois, Urbana-Champaign
John Nicholas Griffis, University of Southern Mississippi
Cameron Gundersen, University of California, Los Angeles
Kenneth Halanych, Auburn University
E. William Hamilton, Washington and Lee University
Monika Havelka, University of Toronto at Mississauga
Tyson Hedrick, University of North Carolina, Chapel Hill
Susan Hengeveld, Indiana University, Bloomington
Albert Herrera, University of Southern California

Kendra Hill, South Dakota State University
 Richard W. Hill, Michigan State University
 Erec B. Hillis, University of California, Berkeley
 Jonathan D. Hillis, Carleton College
 William Huddleston, University of Calgary
 Dianne B. Jennings, Virginia Commonwealth University
 Norman A. Johnson, University of Massachusetts, Amherst
 William H. Karasov, University of Wisconsin, Madison
 Susan Keen, University of California, Davis
 Cornelis Klok, Arizona State University, Tempe
 Olga Ruiz Kopp, Utah Valley University
 William Kroll, Loyola University, Chicago
 Allen Kurta, Eastern Michigan University
 Rebecca Lamb, Ohio State University
 Brenda Leady, University of Toledo
 Hugh Lefcort, Gonzaga University
 Sean C. Lema, University of North Carolina, Wilmington
 Nathan Lents, John Jay College, City University of New York
 Rachel A. Levin, Amherst College
 Donald Levin, University of Texas, Austin
 Bernard Lohr, University of Maryland, Baltimore County
 Barbara Lom, Davidson College
 David J. Longstreth, Louisiana State University
 Catherine Loudon, University of California, Irvine
 Francois Lutzoni, Duke University
 Charles H. Mallory, University of Miami
 Kathi Malueg, University of Colorado, Colorado Springs
 Richard McCarty, Johns Hopkins University
 Sheila McCormick, University of California, Berkeley
 Francis Monette, Boston University
 Leonie Moyle, Indiana University, Bloomington
 Jennifer C. Nauen, University of Delaware

Jacalyn Newman, University of Pittsburgh
 Alexey Nikitin, Grand Valley State University
 Shawn E. Nordell, Saint Louis University
 Tricia Paramore, Hutchinson Community College
 Nancy J. Pelaez, Purdue University
 Robert T. Pennock, Michigan State University
 Roger Persell, Hunter College
 Debra Pires, University of California, Los Angeles
 Crima Pogge, City College of San Francisco
 Jaimie S. Powell, Portland State University
 Susan Richardson, Florida Atlantic University
 David M. Rizzo, University of California, Davis
 Benjamin Rowley, University of Central Arkansas
 Brian Rude, Mississippi State University
 Ann Rushing, Baylor University
 Christina Russin, Northwestern University
 Udo Savalli, Arizona State University, West
 Frieder Schoeck, McGill University
 Paul J. Schulte, University of Nevada, Las Vegas
 Stephen Secor, University of Alabama
 Vijayasaradhi Setaluri, University of Wisconsin, Madison
 H. Bradley Shaffer, University of California, Davis
 Robin Sherman, Nova Southeastern University
 Richard Shingles, Johns Hopkins University
 James Shinkle, Trinity University
 Richard M. Showman, University of South Carolina
 Felisa A. Smith, University of New Mexico
 Ann Berry Somers, University of North Carolina, Greensboro
 Ursula Stochaj, McGill University
 Ken Sweat, Arizona State University, West
 Robin Taylor, Ohio State University
 William Taylor, University of Toledo
 Mark Thogerson, Grand Valley State University

Sharon Thoma, University of Wisconsin, Madison
 Lars Tomanek, California Polytechnic State University
 James Traniello, Boston University
 Jeffrey Travis, State University of New York, Albany
 Terry Trier, Grand Valley State University
 John True, State University of New York, Stony Brook
 Elizabeth Van Volkenburgh, University of Washington
 John Vaughan, St. Petersburg College
 Sara Via, University of Maryland
 Suzanne Wakim, Butte College (Glenn Community College District)
 Randall Walikonis, University of Connecticut
 Cindy White, University of Northern Colorado
 Elizabeth Willott, University of Arizona
 Mark Wilson, Humboldt State University
 Stuart Wooley, California State University, Stanislaus
 Lan Xu, South Dakota State University
 Heping Zhou, Seton Hall University

Accuracy Reviewers

John Alcock, Arizona State University
 Gladys Alexandre, University of Tennessee, Knoxville
 Lawrence A. Alice, Western Kentucky University
 David R. Angelini, American University
 Fabia U. Battistuzzi, Arizona State University
 Arlene Billock, University of Louisiana, Lafayette
 Mary A. Bisson, State University of New York, Buffalo
 Meredith M. Blackwell, Louisiana State University
 Nancy Boury, Iowa State University
 Eldon J. Braun, University of Arizona
 Daniel R. Brooks, University of Toronto
 Jennifer L. Campbell, North Carolina State University
 Peter C. Chabora, Queens College, CUNY
 Patricia Christie, MIT
 Ethan Clotfelter, Amherst College
 Robert Connour, Owens Community College

Peter C. Daniel, Hofstra University
D. Michael Denbow, Virginia Polytechnic Institute
Laura DiCaprio, Ohio State University
Zaldy Doyungan, Texas A&M University, Corpus Christi
Moon Draper, University of Texas, Austin
Richard E. Duhrkopf, Baylor University
Susan A. Dunford, University of Cincinnati
Brent E. Ewers, University of Wyoming
James S. Ferraro, Southern Illinois University
Rachel D. Fink, Mount Holyoke College
John R. Geiser, Western Michigan University
Elizabeth Blinstrup Good, University of Illinois, Urbana-Champaign
Melina E. Hale, University of Chicago
Patricia M. Halpin, University of California, Los Angeles
Jean C. Hardwick, Ithaca College
Monika Havelka, University of Toronto at Mississauga
Frank Healy, Trinity University
Marshal Hedin, San Diego State University
Albert Herrera, University of Southern California
David S. Hibbett, Clark University
James F. Holden, University of Massachusetts, Amherst
Margaret L. Horton, University of North Carolina, Greensboro
Helen Hull-Sanders, Canisius College

C. Darrin Hulsey, University of Tennessee, Knoxville
Timothy Y. James, University of Michigan
Dianne B. Jennings, Virginia Commonwealth University
Norman A. Johnson, University of Massachusetts, Amherst
Susan Jorstad, University of Arizona
Ellen S. Lamb, University of North Carolina, Greensboro
Dennis V. Lavrov, Iowa State University
Hugh Lefort, Gonzaga University
Rachel A. Levin, Amherst College
Bernard Lohr, University of Maryland, Baltimore County
Sharon E. Lynn, College of Wooster
Sarah Mathews, Harvard University
Susan L. Meacham, University of Nevada, Las Vegas
Mona C. Mehdy, University of Texas, Austin
Bradley G. Mehrtens, University of Illinois, Urbana-Champaign
James D. Metzger, Ohio State University
Thomas W. Moon, University of Ottawa
Thomas M. Niesen, San Francisco State University
Theresa O'Halloran, University of Texas, Austin
Thomas L. Pannabecker, University of Arizona
Nancy J. Pelaez, Purdue University

Nicola J. R. Plowes, Arizona State University
Gregory S. Pryor, Francis Marion University
Laurel B. Roberts, University of Pittsburgh
Anjana Sharma, Western Carolina University
Richard M. Showman, University of South Carolina
John B. Skillman, California State University, San Bernardino
John J. Stachowicz, University of California, Davis
Brook O. Swanson, Gonzaga University
Robin A. J. Taylor, Ohio State University
William Taylor, University of Toledo
Steven M. Theg, University of California, Davis
Mark Thogerson, Grand Valley State University
Christopher D. Todd, University of Saskatchewan
Jeffrey Travis, State University of New York, Albany
Joseph S. Walsh, Northwestern University
Andrea Ward, Adelphi University
Barry Williams, Michigan State University
Kenneth Wilson, University of Saskatchewan
Carol L. Wymer, Morehead State University

LIFE's Media and Supplements

BIOPORTAL featuring Prep-U

yourBioPortal.com

BioPortal is the new gateway to all of *Life*'s state-of-the-art online resources for students and instructors. BioPortal includes the breakthrough quizzing engine, Prep-U; a fully interactive eBook; and additional premium learning media. The textbook is tightly integrated with BioPortal via in-text references that connect the printed text and media resources. The result is a powerful, easily-managed online course environment. BioPortal includes the following features and resources:

Life, Ninth Edition eBook

- Integration of all activities, animated tutorials, and other media resources.
- Quick, intuitive navigation to any section or subsection, as well as any printed book page number.
- In-text links to all glossary entries.
- Easy text highlighting.
- A bookmarking feature that allows for quick reference to any page.
- A powerful Notes feature that allows students to add notes to any page.
- A full glossary and index.
- Full-text search, including an additional option to search the glossary and index.
- Automatic saving of all notes, highlighting, and bookmarks.

Additional eBook features for instructors:

- Content Customization: Instructors can easily add pages of their own content and/or hide chapters or sections that they do not cover in their course.
- Instructor Notes: Instructors can choose to create an annotated version of the eBook with their own notes on any page. When students in the course log in, they see the instructor's personalized version of the eBook. Instructor notes can include text, Web links, images, links to all BioPortal content, and more.



Built by educators, Prep-U focuses student study time exactly where it should be, through the use of personalized, adaptive quizzes that move students toward a better grasp of the material—and better grades. For *Life*, Ninth Edition, Prep-U is fully integrated into BioPortal, making it easy for instructors to take advantage of this powerful quizzing engine in their course. Features include:

- Adaptive quizzing
- Automatic results reporting into the BioPortal gradebook

- Misconception index
- Comparison to national data

Student Resources

Diagnostic Quizzing. The diagnostic quiz for each chapter of *Life* assesses student understanding of that chapter, and generates a Personalized Study Plan to effectively focus student study time. The plan includes links to specific textbook sections, animated tutorials, and activities.

Interactive Summaries. For each chapter, these dynamic summaries combine a review of important concepts with links to all of the key figures from the chapter as well as all of the relevant animated tutorials, activities, and key terms.

Animated Tutorials. Over 100 in-depth animated tutorials, in a new format for the Ninth Edition, present complex topics in a clear, easy-to-follow format that combines a detailed animation with an introduction, conclusion, and quiz.

Activities. Over 120 interactive activities help students learn important facts and concepts through a wide range of exercises, such as labeling steps in processes or parts of structures, building diagrams, and identifying different types of organisms.

NEW! Interactive Tutorials. New for the Ninth Edition, these tutorial modules help students master key concepts through hands-on activities that allow them to learn through action. With these tutorials, students can solve problem scenarios by applying concepts from the text, by working with experimental techniques, and by using interactive models to discover how biological mechanisms work. Each tutorial includes a self-assessment quiz that can be assigned.

Interactive Quizzes. Each question includes an image from the textbook, thorough feedback on both correct and incorrect answer choices, references to textbook pages, and links to eBook pages, for quick review.

BioNews from Scientific American. BioNews makes it easy for instructors to bring the dynamic nature of the biological sciences and up-to-the-minute currency into their course. Accessible from within BioPortal, BioNews is a continuously updated feed of current news, podcasts, magazine articles, science blog entries, “strange but true” stories, and more.

NEW! BioNavigator. This unique visual resource is an innovative way to access the wide variety of *Life* media resources. Starting from the whole-Earth view, instructors and students can zoom to any level of biological inquiry, encountering links to a wealth of animations, activities, and tutorials on the full range of topics along the way.

Working with Data. Built around some of the original experiments depicted in the Investigating Life figures, these exercises help build quantitative skills and encourage student in-

terest in how scientists do research, by looking at real experimental data and answering questions based on those data.

Flashcards. For each chapter of the book, there is a set of flashcards that allows the student to review all the key terminology from the chapter. Students can review the terms in study mode, and then quiz themselves on a list of terms.

Experiment Links. For each Investigating Life figure in the textbook, BioPortal includes an overview of the experiment featured in the figure and related research or applications that followed, a link to the original paper, and links to additional information related to the experiment.

Key Terms. The key terminology introduced in each chapter is listed, with definitions and audio pronunciations from the glossary.

Suggested Readings. For each chapter of the book, a list of suggested readings is provided as a resource for further study.

Glossary. The language of biology is often difficult for students taking introductory biology to master, so BioPortal includes a full glossary that features audio pronunciations of all terms.

Statistics Primer. This brief introduction to the use of statistics in biological research explains why statistics are integral to biology, and how some of the most common statistical methods and techniques are used by biologists in their work.

Math for Life. A collection of mathematical shortcuts and references to help students with the quantitative skills they need in the laboratory.

Survival Skills. A guide to more effective study habits. Topics include time management, note-taking, effective highlighting, and exam preparation.

Instructor Resources

Assessment

- Diagnostic Quizzing provides instant class comprehension feedback to instructors, along with targeted lecture resources for those areas requiring the most attention.
- Question banks include questions ranked according to Bloom's taxonomy.
- Question filtering: Allows instructors to select questions based on Bloom's category and/or textbook section.
- Easy-to-use customized assessment tools allow instructors to quickly create quizzes and many other types of assignments using any combination of the questions and resources provided along with their own materials.
- Comprehensive question banks include questions from the test bank, study guide, textbook self-quizzes, and diagnostic quizzes.

Media Resources (see Instructor's Media Library below for details)

- Videos
- PowerPoint® Presentations (Textbook Figures, Lectures, Layered Art)

- Supplemental Photos
- Clicker Questions
- Instructor's Manual
- Lecture Notes

Course Management

- Complete course customization capabilities
- Custom resources/document posting
- Robust Gradebook
- Communication Tools: Announcements, Calendar, Course Email, Discussion Boards

Note: The printed textbook, the eBook, BioPortal, and Prep-U can all be purchased individually as stand-alone items, in addition to being available in a package with the printed textbook.

Student Supplements

Study Guide (ISBN 978-1-4292-3569-3)

Jacalyn Newman, *University of Pittsburgh*; Edward M. Dzialowski, *University of North Texas*; Betty McGuire, *Cornell University*; Lindsay Goodloe, *Cornell University*; and Nancy Guild, *University of Colorado*

For each chapter of the textbook, the *Life* Study Guide offers a variety of study and review tools. The contents of each chapter are broken down into both a detailed review of the Important Concepts covered and a boiled-down Big Picture snapshot. New for the Ninth Edition, Diagram Exercises help students synthesize what they have learned in the chapter through exercises such as ordering concepts, drawing graphs, linking steps in processes, and labeling diagrams. In addition, Common Problem Areas and Study Strategies are highlighted. A set of study questions (both multiple-choice and short-answer) allows students to test their comprehension. All questions include answers and explanations.

Lecture Notebook (ISBN 978-1-4292-3583-9)

This invaluable printed resource consists of all the artwork from the textbook (more than 1,000 images with labels) presented in the order in which they appear in the text, with ample space for note-taking. Because the Notebook has already done the drawing, students can focus more of their attention on the concepts. They will absorb the material more efficiently during class, and their notes will be clearer, more accurate, and more useful when they study from them later.

Companion Website www.thelifewire.com

(Also available as a CD, which can be optionally packaged with the textbook.)

For those students who do not have access to BioPortal, the *Life*, Ninth Edition Companion Website is available free of charge (no access code required). The site features a variety of resources, including animations, flashcards, activities, study ideas, help with math and statistics, and more.

CatchUp Math & Stats

Michael Harris, Gordon Taylor, and Jacquelyn Taylor (ISBN 978-1-4292-0557-3)

This primer will help your students quickly brush up on the quantitative skills they need to succeed in biology. Presented in brief, accessible units, the book covers topics such as working with powers, logarithms, using and understanding graphs, calculating standard deviation, preparing a dilution series, choosing the right statistical test, analyzing enzyme kinetics, and many more.

Student Handbook for Writing in Biology, Third Edition

Karen Knisely, *Bucknell University* (ISBN 978-1-4292-3491-7)

This book provides practical advice to students who are learning to write according to the conventions in biology. Using the standards of journal publication as a model, the author provides, in a user-friendly format, specific instructions on: using biology databases to locate references; paraphrasing for improved comprehension; preparing lab reports, scientific papers, posters; preparing oral presentations in PowerPoint®, and more.

Bioethics and the New Embryology: Springboards for Debate

Scott F. Gilbert, Anna Tyler, and Emily Zackin (ISBN 978-0-7167-7345-0)

Our ability to alter the course of human development ranks among the most significant changes in modern science and has brought embryology into the public domain. The question that must be asked is: Even if we can do such things, should we?

BioStats Basics: A Student Handbook

James L. Gould and Grant F. Gould (ISBN 978-0-7167-3416-1)

BioStats Basics provides introductory-level biology students with a practical, accessible introduction to statistical research. Engaging and informal, the book avoids excessive theoretical and mathematical detail, and instead focuses on how core statistical methods are put to work in biology.

Instructor Media & Supplements

Instructor's Media Library

The *Life*, Ninth Edition Instructor's Media Library (available both online via BioPortal and on disc) includes a wide range of electronic resources to help instructors plan their course, present engaging lectures, and effectively assess student comprehension. The Media Library includes the following resources:

Textbook Figures and Tables. Every image and table from the textbook is provided in both JPEG (high- and low-resolution) and PDF formats. Each figure is provided both with and without balloon captions, and large, complex figures are provided in both a whole and split version.

Unlabeled Figures. Every figure is provided in an unlabeled format, useful for student quizzing and custom presentation development.

Supplemental Photos. The supplemental photograph collection contains over 1,500 photographs (in addition to those in the text), giving instructors a wealth of additional imagery to draw upon.

Animations. Over 100 detailed animations, revised and enlarged for the Ninth Edition, all created from the textbook's art program, and viewable in either narrated or step-through mode.

Videos. A collection of over 200 video segments that covers topics across the entire textbook and helps demonstrate the complexity and beauty of life. Includes the Cell Visualization Videos.

PowerPoint® Resources. For each chapter of the textbook, several different PowerPoint presentations are available. These give instructors the flexibility to build presentations in the manner that best suits their needs. Included are:

- Textbook Figures and Tables
- Lecture Presentation
- Figures with Editable Labels
- Layered Art Figures
- Supplemental Photos
- Videos
- Animations

Clicker Questions. A set of questions written specifically to be used with classroom personal response systems, such as the iClicker system, is provided for each chapter. These questions are designed to reinforce concepts, gauge student comprehension, and engage students in active participation.

Chapter Outlines, Lecture Notes, and the complete **Test File** are all available in Microsoft Word® format for easy use in lecture and exam preparation.

Intuitive Browser Interface provides a quick and easy way to preview and access all of the content on the Instructor's Media Library.

Instructor's Resource Kit

The *Life*, Ninth Edition Instructor's Resource Kit includes a wealth of information to help instructors in the planning and teaching of their course. The Kit includes:

Instructor's Manual, featuring (by chapter):

- A "What's New" guide to the Ninth Edition
- Brief chapter overview
- Chapter outline
- Key terms section with all of the boldface terms from the text

Lecture Notes. Detailed notes for each chapter, which can serve as the basis for lectures, including references to figures and media resources.

Media Guide. A visual guide to the extensive media resources available with the Ninth Edition of *Life*. The guide includes thumbnails and descriptions of every video, animation, lecture PowerPoint®, and supplemental photo in the Media Library, all organized by chapter.

Overhead Transparencies

This set includes over 1,000 transparencies—including all of the four-color line art and all of the tables from the text—along with convenient binders. All figures have been formatted and color-enhanced for clear projection in a wide range of conditions. Labels and images have been resized for improved readability.

Test File

Catherine Ueckert, *Northern Arizona University*; Norman Johnson, *University of Massachusetts*; Paul Nolan, *The Citadel*; Nicola Plowes, *Arizona State University*

The Test File offers more than 5,000 questions, covering the full range of topics presented in the textbook. All questions are referenced to textbook sections and page numbers, and are ranked according to Bloom's taxonomy. Each chapter includes a wide range of multiple choice and fill-in-the-blank questions. In addition, each chapter features a set of diagram questions that involve the student in working with illustrations of structures, graphs, steps in processes, and more. The electronic versions of the Test File (within BioPortal, the Instructor's Media Library, and the Computerized Test Bank CD) also include all of the textbook end-of-chapter Self-Quiz questions, all of the BioPortal Diagnostic Quiz questions, and all of the Study Guide multiple-choice questions.

Computerized Test Bank

The entire printed Test File, plus the textbook end-of-chapter Self-Quizzes, the BioPortal Diagnostic Quizzes, and the Study Guide multiple-choice questions are all included in Wimba's easy-to-use Diploma® software. Designed for both novice and advanced users, Diploma enables instructors to quickly and easily create or edit questions, create quizzes or exams with a "drag-and-drop" feature, publish to online courses, and print paper-based assignments.

Course Management System Support

As a service for *Life* adopters using WebCT, Blackboard, or ANGEL for their courses, full electronic course packs are available.



[www.whfreeman.com/facultylounge/
majorsbio](http://www.whfreeman.com/facultylounge/majorsbio)

NEW! The new Faculty Lounge for Majors Biology is the first publisher-provided website for the majors biology community that lets instructors freely communicate and share peer-reviewed lecture and teaching resources. It is continually updated and vetted by majors biology instructors—there is always something new to see. The Faculty Lounge offers convenient access to peer-recommended and vetted resources, including the following categories: Images, News, Videos, Labs, Lecture Resources, and Educational Research.

In addition, the site includes special areas for resources for lab coordinators, resources and updates from the *Scientific Teaching* series of books, and information on biology teaching workshops.

iClicker

Developed for educators by educators, iClicker is a hassle-free radio-frequency classroom response system that makes it easy for instructors to ask questions, record responses, take attendance, and direct students through lectures as active participants. For more information, visit www.iclicker.com.



www.whfreeman.com/labpartner

NEW! LabPartner is a site designed to facilitate the creation of customized lab manuals. Its database contains a wide selection of experiments published by W. H. Freeman and Hayden-McNeil Publishing. Instructors can preview, choose, and re-order labs, interleave their original experiments, add carbonless graph paper and a pocket folder, and customize the cover both inside and out. LabPartner offers a variety of binding types: paperback, spiral, or loose-leaf. Manuals are printed on-demand once W. H. Freeman receives an order from a campus bookstore or school.



The Scientific Teaching Book Series is a collection of practical guides, intended for all science, technology, engineering and mathematics (STEM) faculty who teach undergraduate and graduate students in these disciplines. The purpose of these books is to help faculty become more successful in all aspects of teaching and learning science, including classroom instruction, mentoring students, and professional development. Authored by well-known science educators, the Series provides concise descriptions of best practices and how to implement them in the classroom, the laboratory, or the department. For readers interested in the research results on which these best practices are based, the books also provide a gateway to the key educational literature.

Scientific Teaching

Jo Handelsman, Sarah Miller, and Christine Pfund,
University of Wisconsin-Madison (ISBN 978-1-4292-0188-9)

NEW! Transformations: Approaches to College Science Teaching

A Collection of Articles from CBE Life Sciences Education
Deborah Allen, *University of Delaware*; Kimberly Tanner, *San Francisco State University* (ISBN 978-1-4292-5335-2)

Contents



PART ONE

THE SCIENCE OF LIFE AND ITS CHEMICAL BASIS

1 Studying Life 2

1.1 What Is Biology? 3

- Cells are the basic unit of life 3
- All of life shares a common evolutionary history 5
- Biological information is contained in a genetic language common to all organisms 6
- Cells use nutrients to supply energy and to build new structures 7
- Living organisms regulate their internal environment 7
- Living organisms interact with one another 8
- Discoveries in biology can be generalized 9

1.2 How Is All Life on Earth Related? 9

- Life arose from non-life via chemical evolution 10
- Cellular structure evolved in the common ancestor of life 10
- Photosynthesis changed the course of evolution 10
- Eukaryotic cells evolved from prokaryotes 11
- Multicellularity arose and cells became specialized 11
- Biologists can trace the evolutionary tree of life 11
- The tree of life is predictive 12

1.3 How Do Biologists Investigate Life? 13

- Observation is an important skill 13
- The scientific method combines observation and logic 14
- Good experiments have the potential to falsify hypotheses 14
- Statistical methods are essential scientific tools 15
- Not all forms of inquiry are scientific 15

1.4 How Does Biology Influence Public Policy? 17

2 Small Molecules and the Chemistry of Life 20

2.1 How Does Atomic Structure Explain the Properties of Matter? 21

- An element consists of only one kind of atom 21
- Each element has a different number of protons 22
- The number of neutrons differs among isotopes 23
- The behavior of electrons determines chemical bonding and geometry 23

2.2 How Do Atoms Bond to Form Molecules? 25

- Covalent bonds consist of shared pairs of electrons 25
- Ionic bonds form by electrical attraction 28
- Hydrogen bonds may form within or between molecules with polar covalent bonds 29

Polar and nonpolar substances: Each interacts best with its own kind 29

2.3 How Do Atoms Change Partners in Chemical Reactions? 30

2.4 What Makes Water So Important for Life? 31

- Water has a unique structure and special properties 31
- Water is an excellent solvent—the medium of life 32
- Aqueous solutions may be acidic or basic 33

An Overview and a Preview 35

3 Proteins, Carbohydrates, and Lipids 38

3.1 What Kinds of Molecules Characterize Living Things? 39

- Functional groups give specific properties to biological molecules 39
- Isomers have different arrangements of the same atoms 40
- The structures of macromolecules reflect their functions 40



Most macromolecules are formed by condensation and broken down by hydrolysis 41

3.2 What Are the Chemical Structures and Functions of Proteins? 42

Amino acids are the building blocks of proteins 42

Peptide linkages form the backbone of a protein 44

The primary structure of a protein is its amino acid sequence 44

The secondary structure of a protein requires hydrogen bonding 46

The tertiary structure of a protein is formed by bending and folding 46

The quaternary structure of a protein consists of subunits 47

Shape and surface chemistry contribute to protein function 48

Environmental conditions affect protein structure 48

Molecular chaperones help shape proteins 49

3.3 What Are the Chemical Structures and Functions of Carbohydrates? 49

Monosaccharides are simple sugars 50

Glycosidic linkages bond monosaccharides 50

Polysaccharides store energy and provide structural materials 52

Chemically modified carbohydrates contain additional functional groups 53

3.4 What Are the Chemical Structures and Functions of Lipids? 54

Fats and oils are hydrophobic 54

Phospholipids form biological membranes 55

Lipids have roles in energy conversion, regulation, and protection 55

The DNA base sequence reveals evolutionary relationships 64

Nucleotides have other important roles 64

4.2 How and Where Did the Small Molecules of Life Originate? 65

Experiments disproved spontaneous generation of life 65

Life began in water 65

Life may have come from outside Earth 66

Prebiotic synthesis experiments model the early Earth 67

4.3 How Did the Large Molecules of Life Originate? 69

Chemical evolution may have led to polymerization 69

There are two theories for the emergence of nucleic acids, proteins, and complex chemistry 69

RNA may have been the first biological catalyst 71

4.4 How Did the First Cells Originate? 72

Experiments describe the origin of cells 72

Some ancient cells left a fossil imprint 73

4 Nucleic Acids and the Origin of Life 60

4.1 What Are the Chemical Structures and Functions of Nucleic Acids? 61

Nucleotides are the building blocks of nucleic acids 61

Base pairing occurs in both DNA and RNA 62

DNA carries information and is expressed through RNA 63

The extracellular matrix supports tissue functions in animals 100

5.5 How Did Eukaryotic Cells Originate? 101

Internal membranes and the nuclear envelope probably came from the plasma membrane 101

Some organelles arose by endosymbiosis 101

6 Cell Membranes 105

6.1 What Is the Structure of a Biological Membrane? 106

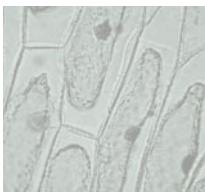
Lipids form the hydrophobic core of the membrane 106

Membrane proteins are asymmetrically distributed 108

Membranes are constantly changing 109

Plasma membrane carbohydrates are recognition sites 110

PART TWO CELLS



5 Cells: The Working Units of Life 76

5.1 What Features Make Cells the Fundamental Units of Life? 77

Cell size is limited by the surface area-to-volume ratio 77

Microscopes reveal the features of cells 79

The plasma membrane forms the outer surface of every cell 79

All cells are classified as either prokaryotic or eukaryotic 80

5.2 What Features Characterize Prokaryotic Cells? 82

Prokaryotic cells share certain features 82

Specialized features are found in some prokaryotes 83

5.3 What Features Characterize Eukaryotic Cells? 84

Compartmentalization is the key to eukaryotic cell function 84

Organelles can be studied by microscopy or isolated for chemical analysis 84

Ribosomes are factories for protein synthesis 85

The nucleus contains most of the genetic information 85

The endomembrane system is a group of interrelated organelles 89

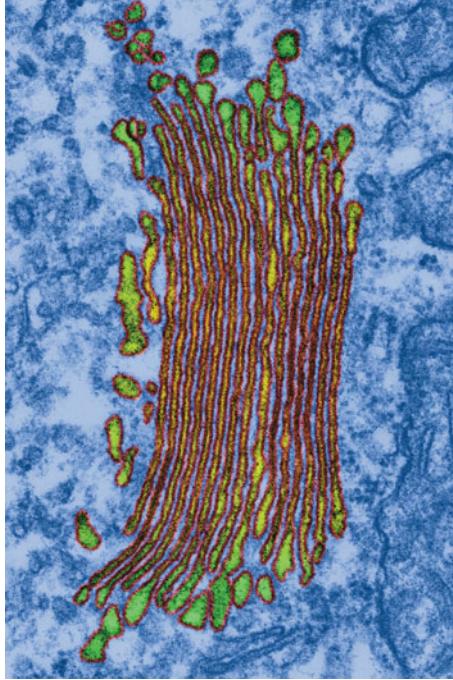
Some organelles transform energy 92

There are several other membrane-enclosed organelles 94

The cytoskeleton is important in cell structure and movement 95

5.4 What Are the Roles of Extracellular Structures? 100

The plant cell wall is an extracellular structure 100



6.2 How Is the Plasma Membrane Involved in Cell Adhesion and Recognition? 111

Cell recognition and cell adhesion involve proteins at the cell surface 111

Three types of cell junctions connect adjacent cells 111

Cell membranes adhere to the extracellular matrix 113

6.3 What Are the Passive Processes of Membrane Transport? 114

Diffusion is the process of random movement toward a state of equilibrium 114

Simple diffusion takes place through the phospholipid bilayer 115

Osmosis is the diffusion of water across membranes 115

Diffusion may be aided by channel proteins 117

Carrier proteins aid diffusion by binding substances 119

6.4 What are the Active Processes of Membrane Transport? 120

Active transport is directional 120

Different energy sources distinguish different active transport systems 121

6.5 How Do Large Molecules Enter and Leave a Cell? 122

Macromolecules and particles enter the cell by endocytosis 123

Receptor-mediated endocytosis is highly specific 123

Exocytosis moves materials out of the cell 124

6.6 What Are Some Other Functions of Membranes? 124

7 Cell Signaling and Communication 128

7.1 What Are Signals, and How Do Cells Respond to Them? 129

Cells receive signals from the physical environment and from other cells 129

A signal transduction pathway involves a signal, a receptor, and responses 130

7.2 How Do Signal Receptors Initiate a Cellular Response? 132

Receptors have specific binding sites for their signals 132

Receptors can be classified by location and function 133

7.3 How Is the Response to a Signal Transduced through the Cell? 136

A protein kinase cascade amplifies a response to ligand binding 136

Second messengers can stimulate protein kinase cascades 137

Second messengers can be derived from lipids 139

Calcium ions are involved in many signal transduction pathways 140

Nitric oxide can act in signal transduction 140

Signal transduction is highly regulated 141

7.4 How Do Cells Change in Response to Signals? 142

Ion channels open in response to signals 142

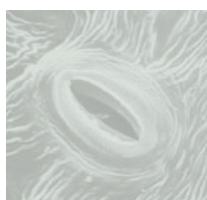
Enzyme activities change in response to signals 143

Signals can initiate DNA transcription 144

7.5 How Do Cells Communicate Directly? 144

Animal cells communicate by gap junctions 144

Plant cells communicate by plasmodesmata 145



PART THREE CELLS AND ENERGY

8 Energy, Enzymes, and Metabolism 148

8.1 What Physical Principles Underlie Biological Energy Transformations? 149

There are two basic types of energy and of metabolism 149

The first law of thermodynamics:

Energy is neither created nor destroyed 150

The second law of thermodynamics:

Disorder tends to increase 150

Chemical reactions release or consume energy 152

Chemical equilibrium and free energy are related 153

8.2 What Is the Role of ATP in Biochemical Energetics? 153

ATP hydrolysis releases energy 154

ATP couples exergonic and endergonic reactions 155

8.3 What Are Enzymes? 156

To speed up a reaction, an energy barrier must be overcome 156

Enzymes bind specific reactants at their active sites 157

Enzymes lower the energy barrier but do not affect equilibrium 157

8.4 How Do Enzymes Work? 158

Enzymes can orient substrates 158

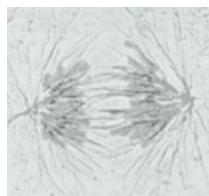
Enzymes can induce strain in the substrate 158

Enzymes can temporarily add chemical groups to substrates 158

Molecular structure determines enzyme function 158

Some enzymes require other molecules in order to function 160	Chemiosmosis is the source of the ATP produced in photophosphorylation 197
The substrate concentration affects the reaction rate 160	
8.5 How Are Enzyme Activities Regulated? 161	10.3 How Is Chemical Energy Used to Synthesize Carbohydrates? 199
Enzymes can be regulated by inhibitors 161	Radioisotope labeling experiments revealed the steps of the Calvin cycle 199
Allosteric enzymes control their activity by changing shape 162	The Calvin cycle is made up of three processes 200
Allosteric effects regulate metabolism 163	Light stimulates the Calvin cycle 201
Enzymes are affected by their environment 164	
9 Pathways that Harvest Chemical Energy 168	10.4 How Do Plants Adapt to the Inefficiencies of Photosynthesis? 202
9.1 How Does Glucose Oxidation Release Chemical Energy? 169	Rubisco catalyzes the reaction of RuBP with O ₂ or CO ₂ 202
Cells trap free energy while metabolizing glucose 169	C ₃ plants undergo photorespiration but C ₄ plants do not 203
Redox reactions transfer electrons and energy 170	CAM plants also use PEP carboxylase 205
The coenzyme NAD ⁺ is a key electron carrier in redox reactions 171	
An overview: Harvesting energy from glucose 171	10.5 How Does Photosynthesis Interact with Other Pathways? 205
9.2 What Are the Aerobic Pathways of Glucose Metabolism? 172	
The energy-investing reactions 1–5 of glycolysis require ATP 174	
The energy-harvesting reactions 6–10 of glycolysis yield NADH and ATP 174	
Pyruvate oxidation links glycolysis and the citric acid cycle 174	
The citric acid cycle completes the oxidation of glucose to CO ₂ 175	
The citric acid cycle is regulated by the concentrations of starting materials 177	
9.3 How Does Oxidative Phosphorylation Form ATP? 177	
The respiratory chain transfers electrons and releases energy 178	
Proton diffusion is coupled to ATP synthesis 178	
9.4 How Is Energy Harvested from Glucose in the Absence of Oxygen? 181	
Cellular respiration yields much more energy than fermentation 183	
The yield of ATP is reduced by the impermeability of some mitochondria to NADH 183	
9.5 How Are Metabolic Pathways Interrelated and Regulated? 184	
Catabolism and anabolism are linked 184	
Catabolism and anabolism are integrated 185	
Metabolic pathways are regulated systems 185	
10 Photosynthesis: Energy from Sunlight 189	
10.1 What Is Photosynthesis? 190	
Experiments with isotopes show that in photosynthesis O ₂ comes from H ₂ O 190	
Photosynthesis involves two pathways 191	
10.2 How Does Photosynthesis Convert Light Energy into Chemical Energy? 192	
Light is a form of energy with dual properties 192	
Molecules become excited when they absorb photons 192	
Absorbed wavelengths correlate with biological activity 193	
Several pigments absorb energy for photosynthesis 194	
Light absorption results in photochemical change 194	
Excited chlorophylls in the reaction center act as electron donors 195	
Reduction leads to electron transport 195	
Noncyclic electron transport produces ATP and NADPH 196	
Cyclic electron transport produces ATP but no NADPH 197	





PART FOUR GENES AND HEREDITY

11 The Cell Cycle and Cell Division 209

11.1 How Do Prokaryotic and Eukaryotic Cells Divide? 210

Prokaryotes divide by binary fission 210
Eukaryotic cells divide by mitosis or meiosis followed by cytokinesis 211

11.2 How Is Eukaryotic Cell Division Controlled? 212

Specific signals trigger events in the cell cycle 213
Growth factors can stimulate cells to divide 215

11.3 What Happens during Mitosis? 215

Prior to mitosis, eukaryotic DNA is packed into very compact chromosomes 215
Overview: Mitosis segregates copies of genetic information 216
The centrosomes determine the plane of cell division 216
The spindle begins to form during prophase 216
Chromosome separation and movement are highly organized 219
Cytokinesis is the division of the cytoplasm 219

11.4 What Role Does Cell Division Play in a Sexual Life Cycle? 221

Asexual reproduction by mitosis results in genetic constancy 221
Sexual reproduction by meiosis results in genetic diversity 221
The number, shapes, and sizes of the metaphase chromosomes constitute the karyotype 223

11.5 What Happens during Meiosis? 224

Meiotic division reduces the chromosome number 225
Chromatid exchanges during meiosis I generate genetic diversity 226
During meiosis homologous chromosomes separate by independent assortment 226

Meiotic errors lead to abnormal chromosome structures and numbers 228
Polyploids have more than two complete sets of chromosomes 229

11.6 In a Living Organism, How Do Cells Die? 229

Cancer cells differ from normal cells 230
Cancer cells lose control over the cell cycle and apoptosis 231
Cancer treatments target the cell cycle 232

12 Inheritance, Genes, and Chromosomes 236

12.1 What Are the Mendelian Laws of Inheritance? 237

Mendel brought new methods to experiments on inheritance 237
Mendel devised a careful research plan 238
Mendel's first experiments involved monohybrid crosses 240
Alleles are different forms of a gene 241
Mendel's first law says that the two copies of a gene segregate 242
Mendel verified his hypothesis by performing a test cross 242
Mendel's second law says that copies of different genes assort independently 244
Punnett squares or probability calculations: A choice of methods 245
Mendel's laws can be observed in human pedigrees 246

12.2 How Do Alleles Interact? 248

New alleles arise by mutation 248
Many genes have multiple alleles 248
Dominance is not always complete 249
In codominance, both alleles at a locus are expressed 249
Some alleles have multiple phenotypic effects 250

12.3 How Do Genes Interact? 250

Hybrid vigor results from new gene combinations and interactions 250

The environment affects gene action 251

Most complex phenotypes are determined by multiple genes and the environment 252

12.4 What Is the Relationship between Genes and Chromosomes? 253

Genes on the same chromosome are linked 253

Genes can be exchanged between chromatids 253

Geneticists can make maps of chromosomes 255

Linkage is revealed by studies of the sex chromosomes 256

Genes on sex chromosomes are inherited in special ways 257

Humans display many sex-linked characters 259

12.5 What Are the Effects of Genes Outside the Nucleus? 259

12.6 How Do Prokaryotes Transmit Genes? 260

Bacteria exchange genes by conjugation 260

Plasmids transfer genes between bacteria 261

13 DNA and Its Role in Heredity 266

13.1 What Is the Evidence that the Gene is DNA? 267

DNA from one type of bacterium genetically transforms another type 267

The transforming principle is DNA 269

Viral replication experiments confirmed that DNA is the genetic material 269

Eukaryotic cells can also be genetically transformed by DNA 270

13.2 What Is the Structure of DNA? 272

The chemical composition of DNA was known 272

Watson and Crick described the double helix 273

Four key features define DNA structure 274

The double-helical structure of DNA is essential to its function 275

13.3 How Is DNA Replicated? 276

Three modes of DNA replication appeared possible 276

An elegant experiment demonstrated that DNA replication is semiconservative 276
 There are two steps in DNA replication 279
 DNA polymerases add nucleotides to the growing chain 279
 Many other proteins assist with DNA polymerization 280
 Telomeres are not fully replicated and are prone to repair 283

13.4 How Are Errors in DNA Repaired? 285

13.5 How Does the Polymerase Chain Reaction Amplify DNA? 286

The polymerase chain reaction makes multiple copies of DNA sequences 286

14 From DNA to Protein: Gene Expression 290

14.1 What Is the Evidence that Genes Code for Proteins? 291

Observations in humans led to the proposal that genes determine enzymes 291

Experiments on bread mold established that genes determine enzymes 292

One gene determines one polypeptide 294

14.2 How Does Information Flow from Genes to Proteins? 294

RNA differs from DNA and plays a vital role in gene expression 294

Two hypotheses were proposed to explain information flow from DNA to protein 295

RNA viruses are exceptions to the central dogma 295

14.3 How Is the Information Content in DNA Transcribed to Produce RNA? 296

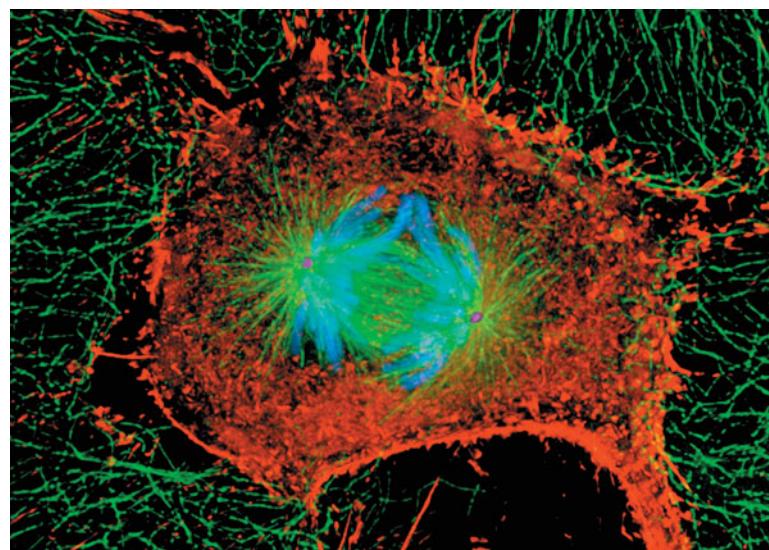
RNA polymerases share common features 296

Transcription occurs in three steps 296

The information for protein synthesis lies in the genetic code 298

14.4 How Is Eukaryotic DNA Transcribed and the RNA Processed? 300

Eukaryotic genes have noncoding sequences 300



Eukaryotic gene transcripts are processed before translation 302

14.5 How Is RNA Translated into Proteins? 304

Transfer RNAs carry specific amino acids and bind to specific codons 304

Activating enzymes link the right tRNAs and amino acids 305

The ribosome is the workbench for translation 306

Translation takes place in three steps 306

Polysome formation increases the rate of protein synthesis 308

14.6 What Happens to Polypeptides after Translation? 310

Signal sequences in proteins direct them to their cellular destinations 310

Many proteins are modified after translation 312

15 Gene Mutation and Molecular Medicine 316

15.1 What Are Mutations? 317

Mutations have different phenotypic effects 317

Point mutations change single nucleotides 318

Chromosomal mutations are extensive changes in the genetic material 320

Mutations can be spontaneous or induced 320

Some base pairs are more vulnerable than others to mutation 322

Mutagens can be natural or artificial 322

Mutations have both benefits and costs 322

15.2 How Are DNA Molecules and Mutations Analyzed? 323

Restriction enzymes cleave DNA at specific sequences 323

Gel electrophoresis separates DNA fragments 324

DNA fingerprinting uses restriction analysis and electrophoresis 325

The DNA barcode project aims to identify all organisms on Earth 326

15.3 How Do Defective Proteins Lead to Diseases? 327

Genetic mutations may make proteins dysfunctional 327

Prion diseases are disorders of protein conformation 329

Most diseases are caused by multiple genes and environment 330

15.4 What DNA Changes Lead to Genetic Diseases? 330

Genetic markers can point the way to important genes 330

Disease-causing mutations may involve any number of base pairs 333

Expanding triplet repeats demonstrate the fragility of some human genes 333

15.5 How Is Genetic Screening Used to Detect Diseases? 334

Screening for disease phenotypes involves analysis of proteins 334

DNA testing is the most accurate way to detect abnormal genes 335

15.6 How Are Genetic Diseases Treated? 337

Genetic diseases can be treated by modifying the phenotype 337
Gene therapy offers the hope of specific treatments 338

16 Regulation of Gene Expression 342

16.1 How Do Viruses Regulate Their Gene Expression? 343

Bacteriophage undergo a lytic cycle 343
Some bacteriophage can carry bacterial genes from one cell to another 345
Some bacteriophage can undergo a lysogenic cycle 345
Eukaryotic viruses have complex regulatory mechanisms 346

16.2 How Is Gene Expression Regulated in Prokaryotes? 348

Regulating gene transcription conserves energy 348
Operons are units of transcriptional regulation in prokaryotes 349
Operator-repressor interactions control transcription in the *lac* and *trp* operons 350
Protein synthesis can be controlled by increasing promoter efficiency 351

16.3 How Is Eukaryotic Gene Transcription Regulated? 352

Transcription factors act at eukaryotic promoters 353
Other proteins can recognize and bind to DNA sequences and regulate transcription 354
Specific protein-DNA interactions underlie binding 354
The expression of sets of genes can be coordinately regulated by transcription factors 355

16.4 How Do Epigenetic Changes Regulate Gene Expression? 356

DNA methylation occurs at promoters and silences transcription 356
Histone protein modifications affect transcription 357
Epigenetic changes induced by the environment can be inherited 358
DNA methylation can result in genomic imprinting 358
Global chromosome changes involve DNA methylation 359

16.5 How Is Eukaryotic Gene Expression Regulated After Transcription? 360

Different mRNAs can be made from the same gene by alternative splicing 360
MicroRNAs are important regulators of gene expression 361
Translation of mRNA can be regulated 361



PART FIVE GENOMES

17 Genomes 365

17.1 How Are Genomes Sequenced? 366

Two approaches were used to sequence the human genome 366
The nucleotide sequence of DNA can be determined 368
High-throughput sequencing has been developed for large genomes 369
Genome sequences yield several kinds of information 370

17.2 What Have We Learned from Sequencing Prokaryotic Genomes? 371

The sequencing of prokaryotic genomes led to new genomics disciplines 371
Some sequences of DNA can move about the genome 372
The sequencing of prokaryotic and viral genomes has many potential benefits 373

Metagenomics allows us to describe new organisms and ecosystems 373
Will defining the genes required for cellular life lead to artificial life? 374

17.3 What Have We Learned from Sequencing Eukaryotic Genomes? 375

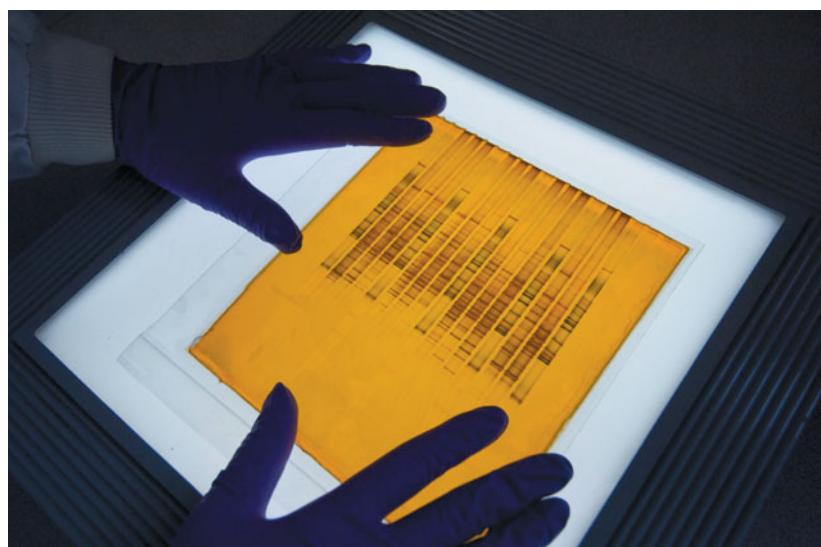
Model organisms reveal many characteristics of eukaryotic genomes 375
Eukaryotes have gene families 377
Eukaryotic genomes contain many repetitive sequences 378

17.4 What Are the Characteristics of the Human Genome? 380

The human genome sequence held some surprises 380
Human genomics has potential benefits in medicine 381

17.5 What Do the New Disciplines Proteomics and Metabolomics Reveal? 382

The proteome is more complex than the genome 382
Metabolomics is the study of chemical phenotype 383



18 Recombinant DNA and Biotechnology 386

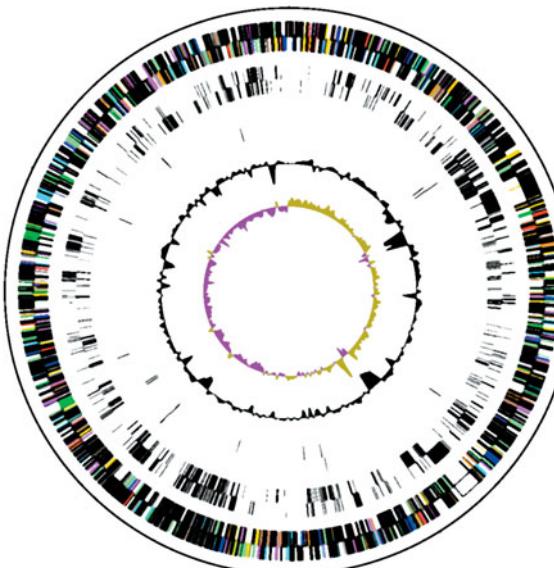
- 18.1 What Is Recombinant DNA? 387
- 18.2 How Are New Genes Inserted into Cells? 389
Genes can be inserted into prokaryotic or eukaryotic cells 389
- Recombinant DNA enters host cells in a variety of ways 389
- Reporter genes identify host cells containing recombinant DNA 390
- 18.3 What Sources of DNA Are Used in Cloning? 392
Libraries provide collections of DNA fragments 392
- cDNA libraries are constructed from mRNA transcripts 392
- Synthetic DNA can be made by PCR or by organic chemistry 393
- DNA mutations can be created in the laboratory 393
- 18.4 What Other Tools Are Used to Study DNA Function? 393
Genes can be inactivated by homologous recombination 393
- Complementary RNA can prevent the expression of specific genes 394
- DNA microarrays can reveal RNA expression patterns 395
- 18.5 What Is Biotechnology? 397
Expression vectors can turn cells into protein factories 397
- 18.6 How Is Biotechnology Changing Medicine, Agriculture, and the Environment? 398
Medically useful proteins can be made by biotechnology 398
- DNA manipulation is changing agriculture 400
- Biotechnology can be used for environmental cleanup 402
- There is public concern about biotechnology 402

19 Differential Gene Expression in Development 405

- 19.1 What Are the Processes of Development? 406
Development involves distinct but overlapping processes 406
- Cell fates become progressively more restricted during development 407
- 19.2 Is Cell Differentiation Irreversible? 408
Plant cells can be totipotent 408
- Nuclear transfer allows the cloning of animals 409
- Multipotent stem cells differentiate in response to environmental signals 410
- Pluripotent stem cells can be obtained in two ways 411
- 19.3 What Is the Role of Gene Expression in Cell Differentiation? 412
Differential gene transcription is a hallmark of cell differentiation 412
- 19.4 How Is Cell Fate Determined? 413
Cytoplasmic segregation can determine polarity and cell fate 413
- Inducers passing from one cell to another can determine cell fates 414
- 19.5 How Does Gene Expression Determine Pattern Formation? 417
Multiple genes interact to determine developmental programmed cell death 417
- Plants have organ identity genes 418
- Morphogen gradients provide positional information 419
- A cascade of transcription factors establishes body segmentation in the fruit fly 420
- Hox genes encode transcription factors 423

20 Development and Evolutionary Change 426

- 20.1 What Is Evo-Devo? 427
Developmental genes in distantly related organisms are similar 427
- 20.2 How Can Mutations With Large Effects Change Only One Part of the Body? 429
Genetic switches govern how the genetic toolkit is used 429
- Modularity allows for differences in the timing and spatial pattern of gene expression 430
- 20.3 How Can Differences among Species Evolve? 432
- 20.4 How Does the Environment Modulate Development? 433
Temperature can determine sex 433
- Organisms use information that predicts future conditions 434
- A variety of environmental signals influence development 435
- 20.5 How Do Developmental Genes Constrain Evolution? 436
Evolution proceeds by changing what's already there 436
- Conserved developmental genes can lead to parallel evolution 436





PART SIX THE PATTERNS AND PROCESSES OF EVOLUTION

21 Evidence and Mechanisms of Evolution 440

21.1 What Facts Form the Basis of Our Understanding of Evolution? 441

Charles Darwin articulated the principle of natural selection 442
Adaptation has two meanings 444
Population genetics provides an underpinning for Darwin's theory 444
Most populations are genetically variable 445
Evolutionary change can be measured by allele and genotype frequencies 446
The genetic structure of a population changes over time, unless certain restrictive conditions exist 447
Deviations from Hardy–Weinberg equilibrium show that evolution is occurring 448

21.2 What Are the Mechanisms of Evolutionary Change? 448

Mutations generate genetic variation 448
Gene flow may change allele frequencies 449
Genetic drift may cause large changes in small populations 449
Nonrandom mating can change genotype frequencies 451

21.3 How Does Natural Selection Result in Evolution? 451

Natural selection can change or stabilize populations 452
Sexual selection influences reproductive success 453

21.4 How Is Genetic Variation Maintained within Populations? 456

Neutral mutations accumulate in populations 456
Sexual recombination amplifies the number of possible genotypes 456
Frequency-dependent selection maintains genetic variation within populations 457

Heterozygote advantage maintains polymorphic loci 457
Much genetic variation in species is maintained in geographically distinct populations 457
21.5 What Are the Constraints on Evolution? 459
Developmental processes constrain evolution 459
Trade-offs constrain evolution 459
Short-term and long-term evolutionary outcomes sometimes differ 460

22 Reconstructing and Using Phylogenies 464

22.1 What Is Phylogeny? 465

All of life is connected through evolutionary history 466
Comparisons among species require an evolutionary perspective 467

22.2 How Are Phylogenetic Trees Constructed? 468

Parsimony provides the simplest explanation for phylogenetic data 470
Phylogenies are reconstructed from many sources of data 470
Mathematical models expand the power of phylogenetic reconstruction 471
The accuracy of phylogenetic methods can be tested 471

22.3 How Do Biologists Use Phylogenetic Trees? 473

Phylogenies help us reconstruct the past 473
Phylogenies allow us to compare and contrast living organisms 474
Ancestral states can be reconstructed 475
Molecular clocks help date evolutionary events 475

22.4 How Does Phylogeny Relate to Classification? 476

Evolutionary history is the basis for modern biological classification 477
Several codes of biological nomenclature govern the use of scientific names 478

23 Species and Their Formation 481

23.1 What Are Species? 482

We can recognize many species by their appearance 482
Species are reproductively isolated lineages on the tree of life 483

23.2 How Do New Species Arise? 484

Gene incompatibilities can produce reproductive isolation in two daughter species 484
Reproductive isolation develops with increasing genetic divergence 485
Geographic barriers give rise to allopatric speciation 485
Sympatric speciation occurs without physical barriers 486

23.3 What Happens When Newly Formed Species Come Together? 489

Prezygotic barriers prevent fertilization 489
Postzygotic barriers can isolate species after fertilization 491
Hybrid zones may form if reproductive isolation is incomplete 492

23.4 Why Do Rates of Speciation Vary? 493

24 Evolution of Genes and Genomes 498

24.1 How Are Genomes Used to Study Evolution? 499

Evolution of genomes results in biological diversity 499
Genes and proteins are compared through sequence alignment 500
Models of sequence evolution are used to calculate evolutionary divergence 501

Experimental studies examine molecular evolution directly 502

24.2 What Do Genomes Reveal About Evolutionary Processes? 505

Much of evolution is neutral 506

Positive and purifying selection can be detected in the genome 506
Genome size and organization also evolve 507

24.3 How Do Genomes Gain and Maintain Functions? 509

Lateral gene transfer can result in the gain of new functions 509
Most new functions arise following gene duplication 510
Some gene families evolve through concerted evolution 511

24.4 What Are Some Applications of Molecular Evolution? 512

Molecular sequence data are used to determine the evolutionary history of genes 512
Gene evolution is used to study protein function 513
In vitro evolution produces new molecules 514
Molecular evolution is used to study and combat diseases 514



Radioisotope dating methods have been expanded and refined 521

25.2 How Have Earth's Continents and Climates Changed over Time? 521

Oxygen concentrations in Earth's atmosphere have changed over time 523
Earth's climate has shifted between hot/humid and cold/dry conditions 524
Volcanoes have occasionally changed the history of life 525
Extraterrestrial events have triggered changes on Earth 525

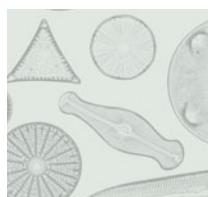
25.3 What Are the Major Events in Life's History? 526

Several processes contribute to the paucity of fossils 526
Precambrian life was small and aquatic 527
Life expanded rapidly during the Cambrian period 527
Many groups of organisms that arose during the Cambrian later diversified 528
Geographic differentiation increased during the Mesozoic era 532
Modern biota evolved during the Cenozoic era 533
The tree of life is used to reconstruct evolutionary events 533

25 The History of Life on Earth 518

25.1 How Do Scientists Date Ancient Events? 519

Radioisotopes provide a way to date rocks 520



PART SEVEN THE EVOLUTION OF DIVERSITY

26 Bacteria and Archaea: The Prokaryotic Domains 536

26.1 How Did the Living World Begin to Diversify? 537

The three domains differ in significant ways 537

26.2 What Are Some Keys to the Success of Prokaryotes? 539

Prokaryotes generally form complex communities 539

Prokaryotes have distinctive cell walls 541

Prokaryotes have distinctive modes of locomotion 541

Prokaryotes reproduce asexually, but genetic recombination can occur 542

Prokaryotes can communicate 542

Prokaryotes have amazingly diverse metabolic pathways 543

26.3 How Can We Resolve Prokaryote Phylogeny? 545

The small size of prokaryotes has hindered our study of their phylogeny 545

The nucleotide sequences of prokaryotes reveal their evolutionary relationships 545

Lateral gene transfer can lead to discordant gene trees 545

The great majority of prokaryote species have never been studied 546

26.4 What Are the Major Known Groups of Prokaryotes? 547

Spirochetes move by means of axial filaments 547

Chlamydias are extremely small parasites 548

Some high-GC Gram-positives are valuable sources of antibiotics 548

Cyanobacteria are important photoautotrophs 548

The low-GC Gram-positives include the smallest cellular organisms 549

The proteobacteria are a large and diverse group 550

Archaea differ in several important ways from bacteria 551

Most Crenarchaeota live in hot and/or acidic places 552
 Euryarchaeota are found in surprising places 552
 Korarchaeota and Nanoarchaeota are less well known 553

26.5 How Do Prokaryotes Affect Their Environments? 553

Prokaryotes are important players in element cycling 553
 Prokaryotes live on and in other organisms 554
 A small minority of bacteria are pathogens 554

26.6 Where Do Viruses Fit into the Tree of Life? 555

Many RNA viruses probably represent escaped genomic components 555
 Some DNA viruses may have evolved from reduced cellular organisms 557

27 The Origin and Diversification of Eukaryotes 560

27.1 How Did the Eukaryotic Cell Arise? 561

The diversity of protists is reflected in both morphology and phylogeny 561
 Cellular features support the monophyly of eukaryotes 561
 The modern eukaryotic cell arose in several steps 564
 Chloroplasts are a study in endosymbiosis 565
 Lateral gene transfer accounts for the presence of some prokaryotic genes in eukaryotes 566

27.2 What Features Account for Protist Diversity? 566

Protists occupy many different niches 566
 Protists have diverse means of locomotion 567
 Protists employ vacuoles in several ways 567
 The cell surfaces of protists are diverse 568

27.3 How Do Protists Affect the World Around Them? 569

Some protists are endosymbionts 569
 Some microbial protists are deadly 570
 We continue to rely on the products of ancient marine protists 571

27.4 How Do Protists Reproduce? 572

Some protists have reproduction without sex, and sex without reproduction 572
 Some protist life cycles feature alternation of generations 573
 Chlorophytes provide examples of several life cycles 574
 The life cycles of some protists require more than one host species 575

27.5 What Are the Evolutionary Relationships among Eukaryotes? 575

Alveolates have sacs under their plasma membrane 575
 Stramenopiles have two unequal flagella, one with hairs 577
 Red algae have a distinctive accessory photosynthetic pigment 579
 Chlorophytes, charophytes, and land plants contain chlorophylls *a* and *b* 580
 Diplomonads and parabasalids are excavates that lack mitochondria 581
 Heteroloboseans alternate between amoeboid forms and forms with flagella 581
 Euglenids and kinetoplastids have distinctive mitochondria and flagella 581
 Foraminiferans have created vast limestone deposits 582
 Radiolarians have thin, stiff pseudopods 582
 Amoebozoans use lobe-shaped pseudopods for locomotion 583

28 Plants without Seeds: From Water to Land 588

28.1 How Did the Land Plants Arise? 589

There are ten major groups of land plants 589
 The land plants arose from a green algal clade 590

28.2 How Did Plants Colonize and Thrive on Land? 591

Adaptations to life on land distinguish land plants from green algae 591
 Nonvascular land plants usually live where water is readily available 592
 Life cycles of land plants feature alternation of generations 592

The sporophytes of nonvascular land plants are dependent on gametophytes 592

28.3 What Features Distinguish the Vascular Plants? 594

Vascular tissues transport water and dissolved materials 594
 Vascular plants have been evolving for almost half a billion years 595
 The earliest vascular plants lacked roots and leaves 596
 The vascular plants branched out 596
 Roots may have evolved from branches 596
 Monilophytes and seed plants have true leaves 597
 Heterospory appeared among the vascular plants 597

28.4 What Are the Major Clades of Seedless Plants? 599

Liverworts may be the most ancient surviving plant clade 599
 Water- and sugar-transport mechanisms first emerged in the mosses 600
 Hornworts have distinctive chloroplasts and sporophytes without stalks 600
 Some vascular plants have vascular tissue but not seeds 601
 The lycophytes are sister to the other vascular plants 602
 Horsetails, whisk ferns, and ferns constitute a clade 602

29 The Evolution of Plants 607

29.1 How Did Seed Plants Become Today's Dominant Vegetation? 608

Features of the seed plant life cycle protect gametes and embryos 608
 The seed is a complex, well-protected package 610
 A change in anatomy enabled seed plants to grow to great heights 611

29.2 What Are the Major Groups of Gymnosperms? 612

Conifers have cones but no motile gametes 613

29.3 What Features Contributed to the Success of the Angiosperms? 615

The sexual structures of angiosperms are flowers 616

Flower structure has evolved over time	616
Angiosperms have coevolved with animals	618
The angiosperm life cycle features double fertilization	619
Angiosperms produce fruits	620
Recent analyses have revealed the oldest split among the angiosperms	620

29.4 How Do Plants Support Our World? 622

Seed plants are our primary food source	622
Seed plants have been sources of medicine since ancient times	623



30 Fungi: Recyclers, Pathogens, Parasites, and Plant Partners 626

30.1 What Is a Fungus? 627

Unicellular fungi are known as yeasts	628
The body of a multicellular fungus is composed of hyphae	629
Fungi are in intimate contact with their environment	630
Fungi reproduce both sexually and asexually	630

30.2 How Do Fungi Interact with Other Organisms? 631

Saprobic fungi are critical to the planetary carbon cycle	631
Fungi may engage in parasitic and predatory interactions	632
Some fungi engage in relationships beneficial to both partners	633
Endophytic fungi protect some plants from pathogens, herbivores, and stress	635

30.3 What Variations Exist among Fungal Life Cycles? 635

Alternation of generations is seen among some aquatic chytrids	635
Terrestrial fungi have separate fusion of cytoplasms and nuclei	638
The dikaryotic condition is a synapomorphy of sac fungi and club fungi	638

30.4 How Have Fungi Evolved and Diversified? 638

Microsporidia are highly reduced, parasitic fungi	639
Chytrids are the only fungi with flagella	639

31 Animal Origins and the Evolution of Body Plans 645

31.1 What Characteristics Distinguish the Animals? 646

Animal monophly is supported by gene sequences and morphology	646
A few basic developmental patterns differentiate major animal groups	648

31.2 What Are the Features of Animal Body Plans? 649

Most animals are symmetrical	649
The structure of the body cavity influences movement	649
Segmentation improves control of movement	650
Appendages have many uses	650

31.3 How Do Animals Get Their Food? 651

Filter feeders capture small prey	652
Herbivores eat plants	652
Predators capture and subdue large prey	652
Parasites live in or on other organisms	653
Detrivores live off the remains of other organisms	654

31.4 How Do Life Cycles Differ among Animals? 654

Most animal life cycles have at least one dispersal stage	655
No life cycle can maximize all benefits	655
Parasite life cycles evolve to facilitate dispersal and overcome host defenses	656
Colonial organisms are composed of genetically identical, physiologically integrated individuals	656

31.5 What Are the Major Groups of Animals? 658

Sponges are loosely organized animals	658
Placozoans are abundant but rarely observed	660
Ctenophores are radially symmetrical and diploblastic	660
Cnidarians are specialized carnivores	661

32 Protostome Animals 666

32.1 What Is a Protostome? 667

Cilia-bearing lophophores and trophophores evolved among the lophotrochozoans	668
Ecdysozoans must shed their cuticles	669
Arrow worms retain some ancestral developmental features	670

32.2 What Features Distinguish the Major Groups of Lophotrochozoans? 671

Bryozoans live in colonies	671
----------------------------	-----

Flatworms and rotifers are structurally diverse relatives 672
 Ribbon worms have a long, protrusible feeding organ 673
 Phoronids and brachiopods use lophophores to extract food from the water 674
 Annelids have segmented bodies 674
 Mollusks have undergone a dramatic evolutionary radiation 676

32.3 What Features Distinguish the Major Groups of Ecdysozoans? 679

Several marine groups have relatively few species 679
 Nematodes and their relatives are abundant and diverse 680

32.4 Why Are Arthropods So Diverse? 681

Arthropod relatives have fleshy, unjointed appendages 682
 Jointed appendages first appeared in the trilobites 682
 Myriapods have many legs 683
 Most chelicerates have four pairs of legs 683
 Crustaceans are diverse and abundant 684

Insects are the dominant terrestrial arthropods 686

An Overview of Protostome Evolution 689

33 Deuterostome Animals 692

33.1 What Is a Deuterostome? 693

33.2 What Are the Major Groups of Echinoderms and Hemichordates? 694

Echinoderms have unique structural features 695
 Hemichordates are wormlike marine deuterostomes 697

33.3 What New Features Evolved in the Chordates? 697

Adults of most cephalochordates and urochordates are sessile 698
 A dorsal supporting structure replaces the notochord in vertebrates 699
 The vertebrate body plan can support large, active animals 700

Fins and swim bladders improved stability and control over locomotion 701

33.4 How Did Vertebrates Colonize the Land? 703

Jointed fins enhanced support for fishes 703
 Amphibians adapted to life on land 704
 Amniotes colonized dry environments 706
 Reptiles adapted to life in many habitats 707
 Crocodilians and birds share their ancestry with the dinosaurs 708
 The evolution of feathers allowed birds to fly 708
 Mammals radiated after the extinction of dinosaurs 709
 Most mammals are therians 710

33.5 What Traits Characterize the Primates? 713

Human ancestors evolved bipedal locomotion 715
 Human brains became larger as jaws became smaller 716
 Humans developed complex language and culture 716



PART EIGHT

FLOWERING PLANTS: FORM AND FUNCTION

34 The Plant Body 719

34.1 What Is the Basic Body Plan of Plants? 720

The root system anchors the plant and takes up water and dissolved minerals 721
 The stem supports leaves and flowers 722
 Leaves are the primary sites of photosynthesis 722

34.2 How Does the Cell Wall Support Plant Growth and Form? 723

Cell walls and vacuoles help determine plant form 723
 The structure of cell walls allows plants to grow 723

34.3 How Do Plant Tissues and Organs Originate? 725

The plant body is constructed from three tissue systems 726
 Cells of the xylem transport water and dissolved minerals 728
 Cells of the phloem transport the products of photosynthesis 728

34.4 How Do Meristems Build a Continuously Growing Plant? 728

Plants increase in size through primary and secondary growth 728
 A hierarchy of meristems generates the plant body 729
 Indeterminate primary growth originates in apical meristems 730
 The root apical meristem gives rise to the root cap and the root primary meristems 730
 The products of the root's primary meristems become root tissues 730

The products of the stem's primary meristems become stem tissues 731

Leaves are determinate organs produced by shoot apical meristems 732

Many eudicot stems and roots undergo secondary growth 733

34.5 How Has Domestication Altered Plant Form? 735

35 Transport in Plants 739

35.1 How Do Plants Take Up Water and Solutes? 740

Water potential differences govern the direction of water movement 740
 Aquaporins facilitate the movement of water across membranes 742
 Uptake of mineral ions requires membrane transport proteins 742

Water and ions pass to the xylem by way of the apoplast and symplast 743

35.2 How Are Water and Minerals Transported in the Xylem? 745

Xylem sap is not pumped by living cells 745

Root pressure alone does not account for xylem transport 745

The transpiration-cohesion-tension mechanism accounts for xylem transport 746

A pressure chamber measures tension in the xylem sap 747

35.3 How Do Stomata Control the Loss of Water and the Uptake of CO₂? 748

The guard cells control the size of the stomatal opening 748

35.4 How Are Substances Translocated in the Phloem? 750

The pressure flow model appears to account for translocation in the phloem 751

The pressure flow model has been experimentally tested 752



Nitrogen fixers make all other life possible 763

Nitrogenase catalyzes nitrogen fixation 763

Some plants and bacteria work together to fix nitrogen 764

Legumes and rhizobia communicate using chemical signals 765

Biological nitrogen fixation does not always meet agricultural needs 765

Plants and bacteria participate in the global nitrogen cycle 766

36.5 How Do Carnivorous and Parasitic Plants Obtain a Balanced Diet? 767

Carnivorous plants supplement their mineral nutrition 767

Parasitic plants take advantage of other plants 767

The plant-parasite relationship is similar to plant-fungi and plant-bacteria associations 768

36 Plant Nutrition 755

36.1 How Do Plants Acquire Nutrients? 756

How does a stationary organism find nutrients? 756

36.2 What Mineral Nutrients Do Plants Require? 757

Deficiency symptoms reveal inadequate nutrition 757

Hydroponic experiments identified essential elements 758

36.3 How Does Soil Structure Affect Plants? 759

Soils are complex in structure 759

Soils form through the weathering of rock 760

Soils are the source of plant nutrition 760

Fertilizers and lime are used in agriculture 761

Plants affect soil fertility and pH 761

36.4 How Do Fungi and Bacteria Increase Nutrient Uptake by Plant Roots? 762

Mycorrhizae expand the root system of plants 762

Soil bacteria are essential in getting nitrogen from air to plant cells 762

Seed germination begins with the uptake of water 774

The embryo must mobilize its reserves 774

Several hormones and photoreceptors help regulate plant growth 774

Signal transduction pathways are involved in all stages of plant development 775

Studies of *Arabidopsis thaliana* have increased our understanding of plant signal transduction 775

37.2 What Do Gibberellins Do? 776

Gibberellins are plant hormones 777

Gibberellins have many effects on plant growth and development 777

Gibberellins act by initiating the breakdown of transcriptional repressors 778

37.3 What Does Auxin Do? 779

Auxin transport is polar and requires carrier proteins 781

Auxin transport mediates responses to light and gravity 781

Auxin affects plant growth in several ways 782

At the molecular level, auxin and gibberellins act similarly 784

37.4 What Are the Effects of Cytokinins, Ethylene, and Brassinosteroids? 784

Cytokinins are active from seed to senescence 784

Ethylene is a gaseous hormone that hastens leaf senescence and fruit ripening 786

Brassinosteroids are plant steroid hormones 786

37 Regulation of Plant Growth 771

37.1 How Does Plant Development Proceed? 772

In early development, the seed germinates and forms a growing seedling 773

Environment cues can initiate seed germination 773

Seed dormancy affords adaptive advantages 774

37.5 How Do Photoreceptors Participate in Plant Growth Regulation? 788

- Phototropins, cryptochromes, and zeaxanthin are blue-light receptors 788
- Phytochromes mediate the effects of red and far-red light 789
- Phytochrome stimulates gene transcription 790
- Circadian rhythms are entrained by light reception 790

38 Reproduction in Flowering Plants 794

38.1 How Do Angiosperms Reproduce Sexually? 795

- The flower is an angiosperm's structure for sexual reproduction 795
- Flowering plants have microscopic gametophytes 796
- Pollination in the absence of water is an evolutionary adaptation 798
- Flowering plants prevent inbreeding 798
- A pollen tube delivers sperm cells to the embryo sac 799
- Angiosperms perform double fertilization 799
- Embryos develop within seeds 800
- Seed development is under hormonal control 801
- Fruits assist in seed dispersal 801

38.2 What Determines the Transition from the Vegetative to the Flowering State? 802

- Apical meristems can become inflorescence meristems 803
- A cascade of gene expression leads to flowering 803
- Photoperiodic cues can initiate flowering 804
- Plants vary in their responses to photoperiodic cues 804
- The length of the night is the key photoperiodic cue determining flowering 804

- The flowering stimulus originates in a leaf 805
- Florigen is a small protein 807
- Flowering can be induced by temperature or gibberellin 808
- Some plants do not require an environmental cue to flower 808

38.3 How Do Angiosperms Reproduce Asexually? 809

- Many forms of asexual reproduction exist 809
- Vegetative reproduction has a disadvantage 810
- Vegetative reproduction is important in agriculture 810

- Some secondary metabolites play multiple roles 821
- Plants respond to herbivory with induced defenses 821
- Why don't plants poison themselves? 822
- The plant doesn't always win the arms race 823

39.3 How Do Plants Deal with Climatic Extremes? 823

- Desert plants have special adaptations to dry conditions 823
- In water-saturated soils, oxygen is scarce 825
- Plants can acclimate to drought stress 826
- Plants have ways of coping with temperature extremes 826

39.4 How Do Plants Deal with Salt and Heavy Metals? 828

- Most halophytes accumulate salt 828
- Some plants can tolerate heavy metals 828

39 Plant Responses to Environmental Challenges 814

39.1 How Do Plants Deal with Pathogens? 815

- Mechanical defenses include physical barriers 815
- Plants can seal off infected parts to limit damage 816
- Plant responses to pathogens may be genetically determined 817
- Receptor-elicitor binding evokes the hypersensitive response 817
- Systemic acquired resistance is a form of long-term "immunity" 818
- Plants develop specific immunity to RNA viruses 818

39.2 How Do Plants Deal with Herbivores? 819

- Herbivory increases the growth of some plants 819
- Mechanical defenses against herbivores are widespread 819
- Plants produce chemical defenses against herbivores 820





PART NINE ANIMALS: FORM AND FUNCTION

40 Physiology, Homeostasis, and Temperature Regulation 832

40.1 How Do Multicellular Animals Supply the Needs of Their Cells? 833

An internal environment makes complex multicellular animals possible 833

Physiological systems maintain homeostasis 834

Cells, tissues, organs, and systems are specialized to serve homeostatic needs 835

Organs consist of multiple tissues 837

40.2 How Does Temperature Affect Living Systems? 838

Q_{10} is a measure of temperature sensitivity 838

Animals acclimatize to seasonal temperatures 839

40.3 How Do Animals Alter Their Heat Exchange with the Environment? 839

Endotherms produce heat metabolically 839

Ectotherms and endotherms respond differently to changes in temperature 840

Energy budgets reflect adaptations for regulating body temperature 841

Both ectotherms and endotherms control blood flow to the skin 842

Some fishes elevate body temperature by conserving metabolic heat 843

Some ectotherms regulate heat production 843

40.4 How Do Mammals Regulate Their Body Temperatures? 844

Basal metabolic rates are correlated with body size and environmental temperature 844

Endotherms respond to cold by producing heat and adapt to cold by reducing heat loss 845

Evaporation of water can dissipate heat, but at a cost 846

The mammalian thermostat uses feedback information 846
Fever helps the body fight infections 847
Turning down the thermostat 847

41 Animal Hormones 851

41.1 What Are Hormones and How Do They Work? 852

Chemical signals can act locally or at a distance 853

Hormonal communication has a long evolutionary history 853

Hormones can be divided into three chemical groups 856

Hormone receptors can be membrane-bound or intracellular 856

Hormone action depends on the nature of the target cell and its receptors 857

41.2 How Do the Nervous and Endocrine Systems Interact? 858

The pituitary connects the nervous and endocrine systems 858

The anterior pituitary is controlled by hypothalamic neurohormones 860

Negative feedback loops regulate hormone secretion 861

41.3 What Are the Major Mammalian Endocrine Glands and Hormones? 861

The thyroid gland secretes thyroxine 861

Three hormones regulate blood calcium concentrations 863

PTH lowers blood phosphate levels 864

Insulin and glucagon regulate blood glucose concentrations 864

The adrenal gland is two glands in one 865

Sex steroids are produced by the gonads 867

Melatonin is involved in biological rhythms and photoperiodicity 868

Many chemicals may act as hormones 868

41.4 How Do We Study Mechanisms of Hormone Action? 868

Hormones can be detected and measured with immunoassays 868
A hormone can act through many receptors 869

42 Immunology: Animal Defense Systems 873

42.1 What Are the Major Defense Systems of Animals? 874

Blood and lymph tissues play important roles in defense 875
White blood cells play many defensive roles 875
Immune system proteins bind pathogens or signal other cells 875

42.2 What Are the Characteristics of the Nonspecific Defenses? 877

Barriers and local agents defend the body against invaders 877
Other nonspecific defenses include specialized proteins and cellular processes 878
Inflammation is a coordinated response to infection or injury 878
Inflammation can cause medical problems 879
Cell signaling pathways stimulate the body's defenses 879

42.3 How Does Specific Immunity Develop? 880

Adaptive immunity has four key features 880
Two types of specific immune responses interact: an overview 881
Genetic changes and clonal selection generate the specific immune response 882
Immunity and immunological memory result from clonal selection 883
Vaccines are an application of immunological memory 883
Animals distinguish self from nonself and tolerate their own antigens 884

42.4 What Is the Humoral Immune Response? 885

- Some B cells develop into plasma cells 885
- Different antibodies share a common structure 885
- There are five classes of immunoglobulins 886
- Monoclonal antibodies have many uses 886

42.5 What Is the Cellular Immune Response? 887

- T cell receptors bind to antigens on cell surfaces 888
- MHC proteins present antigen to T cells 888
- T-helper cells and MHC II proteins contribute to the humoral immune response 889
- Cytotoxic T cells and MHC I proteins contribute to the cellular immune response 889
- Regulatory T cells suppress the humoral and cellular immune responses 889
- MHC proteins are important in tissue transplants 891

42.6 How Do Animals Make So Many Different Antibodies? 891

- Antibody diversity results from DNA rearrangement and other mutations 892
- The constant region is involved in immunoglobulin class switching 893

42.7 What Happens When the Immune System Malfunctions? 894

- Allergic reactions result from hypersensitivity 894

- Autoimmune diseases are caused by reactions against self antigens 895
- AIDS is an immune deficiency disorder 895

43 Animal Reproduction 899

43.1 How Do Animals Reproduce without Sex? 900

- Budding and regeneration produce new individuals by mitosis 900
- Parthenogenesis is the development of unfertilized eggs 901

43.2 How Do Animals Reproduce Sexually? 902

- Gametogenesis produces eggs and sperm 902
- Fertilization is the union of sperm and egg 905
- Getting eggs and sperm together 906
- An individual animal can function as both male and female 907
- The evolution of vertebrate reproductive systems parallels the move to land 907
- Animals with internal fertilization are distinguished by where the embryo develops 908

43.3 How Do the Human Male and Female Reproductive Systems Work? 909

- Male sex organs produce and deliver semen 909
- Male sexual function is controlled by hormones 911

- Female sex organs produce eggs, receive sperm, and nurture the embryo 912

The ovarian cycle produces a mature egg 913

The uterine cycle prepares an environment for the fertilized egg 913

Hormones control and coordinate the ovarian and uterine cycles 914

In pregnancy, hormones from the extraembryonic membranes take over 915

Childbirth is triggered by hormonal and mechanical stimuli 916

43.4 How Can Fertility Be Controlled? 916

- Human sexual responses have four phases 916
- Humans use a variety of methods to control fertility 917
- Reproductive technologies help solve problems of infertility 919

44 Animal Development 922

44.1 How Does Fertilization Activate Development? 923

- The sperm and the egg make different contributions to the zygote 923
- Rearrangements of egg cytoplasm set the stage for determination 924
- Cleavage repackages the cytoplasm 925
- Early cell divisions in mammals are unique 926
- Specific blastomeres generate specific tissues and organs 927

44.2 How Does Gastrulation Generate Multiple Tissue Layers? 928

- Invagination of the vegetal pole characterizes gastrulation in the sea urchin 928
- Gastrulation in the frog begins at the gray crescent 929
- The dorsal lip of the blastopore organizes embryo formation 930
- Transcription factors underlie the organizer's actions 931
- The organizer changes its activity as it migrates from the dorsal lip 932
- Reptilian and avian gastrulation is an adaptation to yolked eggs 933
- Placental mammals retain the avian-reptilian gastrulation pattern but lack yolk 934



44.3 How Do Organs and Organ Systems Develop? 935

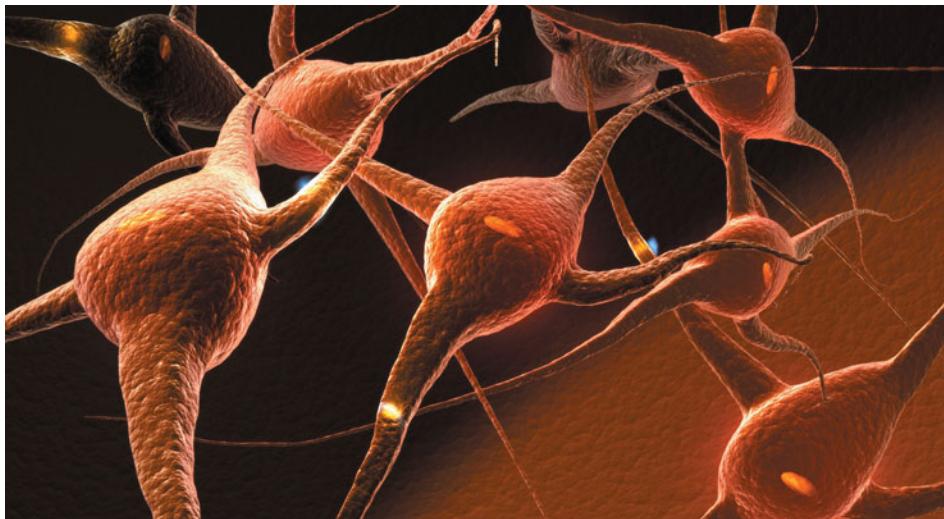
- The stage is set by the dorsal lip of the blastopore 935
- Body segmentation develops during neurulation 935
- Hox genes control development along the anterior-posterior axis 936

44.4 How is the Growing Embryo Sustained? 937

- Extraembryonic membranes form with contributions from all germ layers 937
- Extraembryonic membranes in mammals form the placenta 938

44.5 What Are the Stages of Human Development? 939

- Organ development begins in the first trimester 939
- Organ systems grow and mature during the second and third trimesters 939
- Developmental changes continue throughout life 940



45 Neurons and Nervous Systems 943

45.1 What Cells Are Unique to the Nervous System? 944

- Neural networks range in complexity 944
- Neurons are the functional units of nervous systems 945
- Glia are also important components of nervous systems 947

45.2 How Do Neurons Generate and Transmit Electrical Signals? 948

- Simple electrical concepts underlie neural function 948
- Membrane potentials can be measured with electrodes 948
- Ion transporters and channels generate membrane potentials 948
- Ion channels and their properties can now be studied directly 951
- Gated ion channels alter membrane potential 952
- Graded changes in membrane potential can integrate information 952
- Sudden changes in Na^+ and K^+ channels generate action potentials 953
- Action potentials are conducted along axons without loss of signal 955

Action potentials can jump along axons 955

45.3 How Do Neurons Communicate with Other Cells? 956

- The neuromuscular junction is a model chemical synapse 956
- The arrival of an action potential causes the release of neurotransmitter 956
- Synaptic functions involve many proteins 957
- The postsynaptic membrane responds to neurotransmitter 957
- Synapses between neurons can be excitatory or inhibitory 958
- The postsynaptic cell sums excitatory and inhibitory input 958
- Synapses can be fast or slow 958
- Electrical synapses are fast but do not integrate information well 959
- The action of a neurotransmitter depends on the receptor to which it binds 959
- Glutamate receptors may be involved in learning and memory 960
- To turn off responses, synapses must be cleared of neurotransmitter 960
- The diversity of receptors makes drug specificity possible 961

Sensation depends on which neurons receive action potentials from sensory cells 967

Many receptors adapt to repeated stimulation 967

46.2 How Do Sensory Systems Detect Chemical Stimuli? 967

- Arthropods are good models for studying chemoreception 968
- Olfaction is the sense of smell 968
- The vomeronasal organ contains chemoreceptors 969
- Gustation is the sense of taste 969

46.3 How Do Sensory Systems Detect Mechanical Forces? 970

- Many different cells respond to touch and pressure 970
- Mechanoreceptors are found in muscles, tendons, and ligaments 971
- Auditory systems use hair cells to sense sound waves 972
- Hair cells are sensitive to being bent 974
- Hair cells detect forces of gravity and momentum 975
- Hair cells are evolutionarily conserved 975

46.4 How Do Sensory Systems Detect Light? 976

- Rhodopsins are responsible for photosensitivity 976
- Rod cells respond to light 976
- Invertebrates have a variety of visual systems 978
- Image-forming eyes evolved independently in vertebrates and cephalopods 979
- The vertebrate retina receives and processes visual information 980

46 Sensory Systems 964

46.1 How Do Sensory Cells Convert Stimuli into Action Potentials? 965

- Sensory receptor proteins act on ion channels 965
- Sensory transduction involves changes in membrane potentials 966

47 The Mammalian Nervous System: Structure and Higher Function 985

47.1 How Is the Mammalian Nervous System Organized? 986

- A functional organization of the nervous system is based on flow and type of information 986
- The vertebrate CNS develops from the embryonic neural tube 986
- The spinal cord transmits and processes information 988
- The reticular system alerts the forebrain 989
- The core of the forebrain controls physiological drives, instincts, and emotions 989
- Regions of the telencephalon interact to produce consciousness and control behavior 990
- The human brain is off the curve 992

47.2 How Is Information Processed by Neural Networks? 993

- The autonomic nervous system controls involuntary physiological functions 993
- Patterns of light falling on the retina are integrated by the visual cortex 994
- Cortical cells receive input from both eyes 997

47.3 Can Higher Functions Be Understood in Cellular Terms? 998

- Sleep and dreaming are reflected in electrical patterns in the cerebral cortex 999
- Language abilities are localized in the left cerebral hemisphere 1000
- Some learning and memory can be localized to specific brain areas 1000
- We still cannot answer the question “What is consciousness?” 1002

48 Musculoskeletal Systems 1006

48.1 How Do Muscles Contract? 1007

- Sliding filaments cause skeletal muscle to contract 1007
- Actin-myosin interactions cause filaments to slide 1009
- Actin-myosin interactions are controlled by calcium ions 1010
- Cardiac muscle is similar to and different from skeletal muscle 1012
- Smooth muscle causes slow contractions of many internal organs 1012
- Single skeletal muscle twitches are summed into graded contractions 1014

48.2 What Determines Muscle Performance? 1015

- Muscle fiber types determine endurance and strength 1015

- A muscle has an optimal length for generating maximum tension 1016
- Exercise increases muscle strength and endurance 1017
- Muscle ATP supply limits performance 1017
- Insect muscle has the greatest rate of cycling 1018

48.3 How Do Skeletal Systems and Muscles Work Together? 1018

- A hydrostatic skeleton consists of fluid in a muscular cavity 1018
- Exoskeletons are rigid outer structures 1019
- Vertebrate endoskeletons consist of cartilage and bone 1019
- Bones develop from connective tissues 1020
- Bones that have a common joint can work as a lever 1021

49 Gas Exchange in Animals 1025

49.1 What Physical Factors Govern Respiratory Gas Exchange? 1026

- Diffusion is driven by concentration differences 1026
- Fick's law applies to all systems of gas exchange 1027
- Air is a better respiratory medium than water 1027
- High temperatures create respiratory problems for aquatic animals 1028
- O₂ availability decreases with altitude 1028
- CO₂ is lost by diffusion 1028

49.2 What Adaptations Maximize Respiratory Gas Exchange? 1029

- Respiratory organs have large surface areas 1029
- Transporting gases to and from exchange surfaces optimizes partial pressure gradients 1029
- Insects have airways throughout their bodies 1029
- Fish gills use countercurrent flow to maximize gas exchange 1030
- Birds use unidirectional ventilation to maximize gas exchange 1031
- Tidal ventilation produces dead space that limits gas exchange efficiency 1032



49.3 How Do Human Lungs Work? 1035

Respiratory tract secretions aid ventilation 1035
Lungs are ventilated by pressure changes in the thoracic cavity 1035

49.4 How Does Blood Transport Respiratory Gases? 1037

Hemoglobin combines reversibly with O₂ 1037
Myoglobin holds an O₂ reserve 1038
Hemoglobin's affinity for O₂ is variable 1038
CO₂ is transported as bicarbonate ions in the blood 1039

49.5 How Is Breathing Regulated? 1040

Breathing is controlled in the brainstem 1040
Regulating breathing requires feedback information 1041

50 Circulatory Systems 1045**50.1 Why Do Animals Need a Circulatory System? 1046**

Some animals do not have a circulatory system 1046
Circulatory systems can be open or closed 1046
Open circulatory systems move extracellular fluid 1047
Closed circulatory systems circulate blood through a system of blood vessels 1047

50.2 How Have Vertebrate Circulatory Systems Evolved? 1048

Fishes have a two-chambered heart 1048
Amphibians have a three-chambered heart 1049
Reptiles have exquisite control of pulmonary and systemic circulation 1049
Birds and mammals have fully separated pulmonary and systemic circuits 1050

50.3 How Does the Mammalian Heart Function? 1051

Blood flows from right heart to lungs to left heart to body 1051
The heartbeat originates in the cardiac muscle 1053
A conduction system coordinates the contraction of heart muscle 1054

Electrical properties of ventricular muscles sustain heart contraction 1054

The ECG records the electrical activity of the heart 1055

50.4 What Are the Properties of Blood and Blood Vessels? 1056

Red blood cells transport respiratory gases 1056
Platelets are essential for blood clotting 1057
Blood circulates throughout the body in a system of blood vessels 1058
Materials are exchanged in capillary beds by filtration, osmosis, and diffusion 1058
Blood flows back to the heart through veins 1060
Lymphatic vessels return interstitial fluid to the blood 1061
Vascular disease is a killer 1061

50.5 How Is the Circulatory System Controlled and Regulated? 1062

Autoregulation matches local blood flow to local need 1062
Arterial pressure is regulated by hormonal and neural mechanisms 1063

Tubular guts have an opening at each end 1075

Digestive enzymes break down complex food molecules 1076

51.3 How Does the Vertebrate Gastrointestinal System Function? 1077

The vertebrate gut consists of concentric tissue layers 1077
Mechanical activity moves food through the gut and aids digestion 1078
Chemical digestion begins in the mouth and the stomach 1079
Stomach ulcers can be caused by a bacterium 1079
The stomach gradually releases its contents to the small intestine 1080
Most chemical digestion occurs in the small intestine 1081
Nutrients are absorbed in the small intestine 1082
Absorbed nutrients go to the liver 1083
Water and ions are absorbed in the large intestine 1083
Herbivores rely on microorganisms to digest cellulose 1083

51.4 How Is the Flow of Nutrients Controlled and Regulated? 1084

Hormones control many digestive functions 1085
The liver directs the traffic of the molecules that fuel metabolism 1085
Regulating food intake is important 1087

51 Nutrition, Digestion, and Absorption 1067**51.1 What Do Animals Require from Food? 1068**

Energy needs and expenditures can be measured 1068
Sources of energy can be stored in the body 1070
Food provides carbon skeletons for biosynthesis 1070
Animals need mineral elements for a variety of functions 1071
Animals must obtain vitamins from food 1072
Nutrient deficiencies result in diseases 1073

51.2 How Do Animals Ingest and Digest Food? 1074

The food of herbivores is often low in energy and hard to digest 1074
Carnivores must detect, capture, and kill prey 1074
Vertebrate species have distinctive teeth 1074
Digestion usually begins in a body cavity 1075

52 Salt and Water Balance and Nitrogen Excretion 1091**52.1 How Do Excretory Systems Maintain Homeostasis? 1092**

Water enters or leaves cells by osmosis 1092
Excretory systems control extracellular fluid osmolarity and composition 1093
Animals can be osmoconformers or osmoregulators 1093
Animals can be ionic conformers or ionic regulators 1093

52.2 How Do Animals Excrete Nitrogen? 1094

Animals excrete nitrogen in a number of forms 1094

Most species produce more than one nitrogenous waste 1095

52.3 How Do Invertebrate Excretory Systems Work? 1095

The protonephridia of flatworms excrete water and conserve salts 1095

The metanephridia of annelids process coelomic fluid 1096

The Malpighian tubules of insects depend on active transport 1097

52.4 How Do Vertebrates Maintain Salt and Water Balance? 1097

Marine fishes must conserve water 1098

Terrestrial amphibians and reptiles must avoid desiccation 1098

Mammals can produce highly concentrated urine 1098

The nephron is the functional unit of the vertebrate kidney 1098

Blood is filtered into Bowman's capsule 1099

The renal tubules convert glomerular filtrate to urine 1100

52.5 How Does the Mammalian Kidney Produce Concentrated Urine? 1100

Kidneys produce urine, and the bladder stores it 1100

Nephrons have a regular arrangement in the kidney 1100

Most of the glomerular filtrate is reabsorbed by the proximal convoluted tubule 1101

The loop of Henle creates a concentration gradient in the renal medulla 1102

Water permeability of kidney tubules depends on water channels 1103

The distal convoluted tubule fine-tunes the composition of the urine 1104

Urine is concentrated in the collecting duct 1104

The kidneys help regulate acid-base balance 1104

Kidney failure is treated with dialysis 1105

52.6 How Are Kidney Functions Regulated? 1107

Glomerular filtration rate is regulated 1107

Blood osmolarity and blood pressure are regulated by ADH 1107

The heart produces a hormone that helps lower blood pressure 1109

53 Animal Behavior 1113

53.1 What Are the Origins of Behavioral Biology? 1114

Conditioning was the focus of behaviorists 1114

Fixed action patterns were the focus of ethologists 1115

Ethologists probed the causes of behavior 1116

53.2 How Can Genes Influence Behavior? 1117

Breeding experiments can show whether behavioral phenotypes are genetically determined 1117

Knockout experiments can reveal the roles of specific genes 1117

Behaviors are controlled by gene cascades 1118

53.3 How Does Behavior Develop? 1119

Hormones can determine behavioral potential and timing 1119

Some behaviors can be acquired only at certain times 1120

Bird song learning involves genetics, imprinting, and hormonal timing 1121

The timing and expression of bird song are under hormonal control 1122

53.4 How Does Behavior Evolve? 1123

Animals must make many behavioral choices 1123

Behaviors have costs and benefits 1123

Cost-benefit analysis can be applied to foraging behavior 1125

53.5 What Physiological Mechanisms Underlie Behavior? 1127

Biological rhythms coordinate behavior with environmental cycles 1127

Animals must find their way around their environment 1129

Animals use multiple modalities to communicate 1131

53.6 How Does Social Behavior Evolve? 1133

Mating systems maximize the fitness of both partners 1134

Fitness can include more than producing offspring 1135

Eusociality is the extreme result of kin selection 1136

Group living has benefits and costs 1137





PART TEN ECOLOGY

54 Ecology and the Distribution of Life 1140

54.1 What Is Ecology? 1141

- Ecology is not the same as environmentalism 1141
- Ecologists study biotic and abiotic components of ecosystems 1142

54.2 Why Do Climates Vary Geographically? 1142

- Global climates are solar powered 1142
- Solar energy input determines atmospheric circulation patterns 1142
- Global oceanic circulation is driven by wind patterns 1143
- Organisms adapt to climatic challenges 1144

54.3 What Is a Biome? 1146

- Tundra is found at high latitudes and high elevations 1147
- Evergreen trees dominate boreal forests 1148
- Temperate deciduous forests change with the seasons 1149
- Temperate grasslands are widespread 1150
- Hot deserts form around 30° latitude 1151
- Cold deserts are high and dry 1152
- Chaparral has hot, dry summers and wet, cool winters 1153
- Thorn forests and tropical savannas have similar climates 1154
- Tropical deciduous forests occur in hot lowlands 1155
- Tropical evergreen forests are rich in species 1156
- Biome distribution is not determined solely by temperature 1157

54.4 What Is a Biogeographic Region? 1157

- Geological history influenced the distribution of organisms 1157
- Two scientific advances changed the field of biogeography 1159
- Biotic interchange follows fusion of land masses 1160

- Vicariant events influence distribution patterns 1161
- Humans exert a powerful influence on biogeographic patterns 1162

54.5 How Is Life Distributed in Aquatic Environments? 1163

- The oceans can be divided into several life zones 1163
- Freshwater environments may be rich in species 1164
- Estuaries have characteristics of both freshwater and marine environments 1164

- Different population regulation factors lead to different life histories 1176
- Several factors explain why some species achieve higher population densities than others 1176
- Evolutionary history may explain species abundances 1178

55.4 How Does Habitat Variation Affect Population Dynamics? 1178

- Many populations live in separated habitat patches 1178
- Corridors allow subpopulations to persist 1179

55.5 How Can Populations Be Managed Scientifically? 1180

- Population management plans must take life history strategies into account 1180
- Population management plans must be guided by the principles of population dynamics 1180
- Human population increase has been exponential 1181

55 Population Ecology 1167

55.1 How Do Ecologists Study Populations? 1168

- Ecologists use a variety of approaches to count and track individuals 1168
- Population densities can be estimated from samples 1169
- Populations have age structures and dispersion patterns 1169
- Changes in population size can be estimated from repeated density measurements 1171
- Life tables track demographic events 1171

55.2 How Do Environmental Conditions Affect Life Histories? 1173

- Survivorship and fecundity determine a population's intrinsic rate of increase 1173
- Life history traits vary with environmental conditions 1173
- Life history traits are influenced by interspecific interactions 1174

55.3 What Factors Limit Population Densities? 1174

- All populations have the potential for exponential growth 1174
- Logistic growth occurs as a population approaches its carrying capacity 1175
- Population growth can be limited by density-dependent or density-independent factors 1176



56 Species Interaction and Coevolution 1185

56.1 What Types of Interactions Do Ecologists Study? 1186

- Interactions among species can be grouped into several categories 1186
- Some types of interactions result in coevolution 1188

56.2 How Do Antagonistic Interactions Evolve? 1188

- Predator-prey interactions result in a range of adaptations 1189
- Herbivory is a widespread interaction 1190
- Microparasite-host interactions may be pathogenic 1192
- Most ectoparasites have adaptations for holding onto their hosts 1193

56.3 How Do Mutualistic Interactions Evolve? 1194

- Plants and pollinators exchange food for pollen transport 1194
- Plants and frugivores exchange food for seed transport 1196
- Some mutualistic partners exchange food for care or transport 1196
- Some mutualistic partners exchange food or housing for defense 1196

56.4 What Are the Outcomes of Competition? 1198

- Competition is widespread because all species share resources 1198
- Interference competition may restrict habitat use 1198
- Exploitation competition may lead to coexistence 1198
- Species may compete indirectly for a resource 1199
- Consumers may influence the outcome of competition 1199
- Competition may determine a species' niche 1199



Fewer individuals and less biomass can be supported at higher trophic levels 1206

Productivity and species richness are linked 1206

57.2 How Do Interactions among Species Influence Community Structure? 1208

- Species interactions can cause trophic cascades 1208
- Keystone species have wide-ranging effects 1209

57.3 What Patterns of Species Diversity Have Ecologists Observed? 1210

- A community's diversity can be measured with a diversity index 1210
- Latitudinal gradients in diversity are observed in both hemispheres 1211
- The theory of island biogeography suggests that species richness reaches an equilibrium 1212

57.4 How Do Disturbances Affect Ecological Communities? 1214

- Succession is the predictable pattern of change in a community after a disturbance 1214
- Both facilitation and inhibition influence succession 1215
- Cyclical succession requires adaptation to periodic disturbances 1216
- Heterotrophic succession generates distinctive communities 1216

57.5 How Does Species Richness Influence Community Stability? 1217

Species-rich communities use resources more efficiently 1217

Diversity, productivity, and stability differ between natural and managed communities 1218

58 Ecosystems and Global Ecology 1221

58.1 What Are the Compartments of the Global Ecosystem? 1222

- Energy flows and materials cycle through ecosystems 1222
- The atmosphere regulates temperatures close to Earth's surface 1223
- The oceans receive materials from the other compartments 1224
- Water moves rapidly through lakes and streams 1225
- Land covers about a quarter of Earth's surface 1226

58.2 How Does Energy Flow through the Global Ecosystem? 1227

- The geographic distribution of energy flow is uneven 1227
- Human activities modify the flow of energy 1228

58.3 How Do Materials Cycle through the Global Ecosystem? 1229

- Water transfers materials from one compartment to another 1229
- Fire is a major mover of elements 1230

57 Community Ecology 1203

57.1 What Are Ecological Communities? 1204

- Energy enters communities through primary producers 1204
- Consumers use diverse sources of energy 1205

The carbon cycle has been altered by human activities 1230	Biodiversity has great value to human society 1244	Degraded ecosystems can be restored 1252
Recent perturbations of the nitrogen cycle have had adverse effects on ecosystems 1232	59.2 How Do Biologists Predict Changes in Biodiversity? 1245	Disturbance patterns sometimes need to be restored 1253
The burning of fossil fuels affects the sulfur cycle 1234	Our knowledge of biodiversity is incomplete 1245	Ending trade is crucial to saving some species 1254
The global phosphorus cycle lacks a significant atmospheric component 1235	We can predict the effects of human activities on biodiversity 1246	Invasions of exotic species must be controlled or prevented 1254
Other biogeochemical cycles are also important 1236	59.3 What Factors Threaten Species Persistence? 1247	Biodiversity can have market value 1255
Biogeochemical cycles interact 1236	Species are endangered by the degradation, destruction, and fragmentation of their habitats 1247	Simple changes can help protect biodiversity 1256
58.4 What Services Do Ecosystems Provide? 1237	Overexploitation has driven many species to extinction 1248	Captive breeding programs can maintain a few species 1257
58.5 How Can Ecosystems Be Sustainably Managed? 1238	Invasive predators, competitors, and pathogens threaten many species 1249	The legacy of Samuel Plimsoll 1257
59 Conservation Biology 1242	Rapid climate change can cause species extinctions 1249	
59.1 What Is Conservation Biology? 1243	59.4 What Strategies Do Biologists Use to Protect Biodiversity? 1251	
Conservation biology aims to protect and manage biodiversity 1243	Protected areas preserve habitat and prevent overexploitation 1251	
		Appendix A: The Tree of Life 1260
		Appendix B: Some Measurements Used in Biology 1267
		Answers to Self-Quizzes A-1
		Glossary G-1
		Illustration Credits C-1
		Index I-1

Life

The Science of Biology

NINTH EDITION

Studying Life

Why are frogs croaking?

Amphibians—frogs, toads, and salamanders—have been around for a long time. They watched the dinosaurs come and go. But today amphibian populations around the world are in dramatic decline, with more than a third of the world's amphibian species threatened with extinction. Why?

Biologists work to answer this question by making observations and doing experiments. A number of factors may be involved, and one possible cause may be the effects of agricultural pesticides and herbicides. Several studies have shown that many of these chemicals tested at realistic concentrations do not kill amphibians. But Tyrone Hayes, a biologist at the University of California at Berkeley, probed deeper.

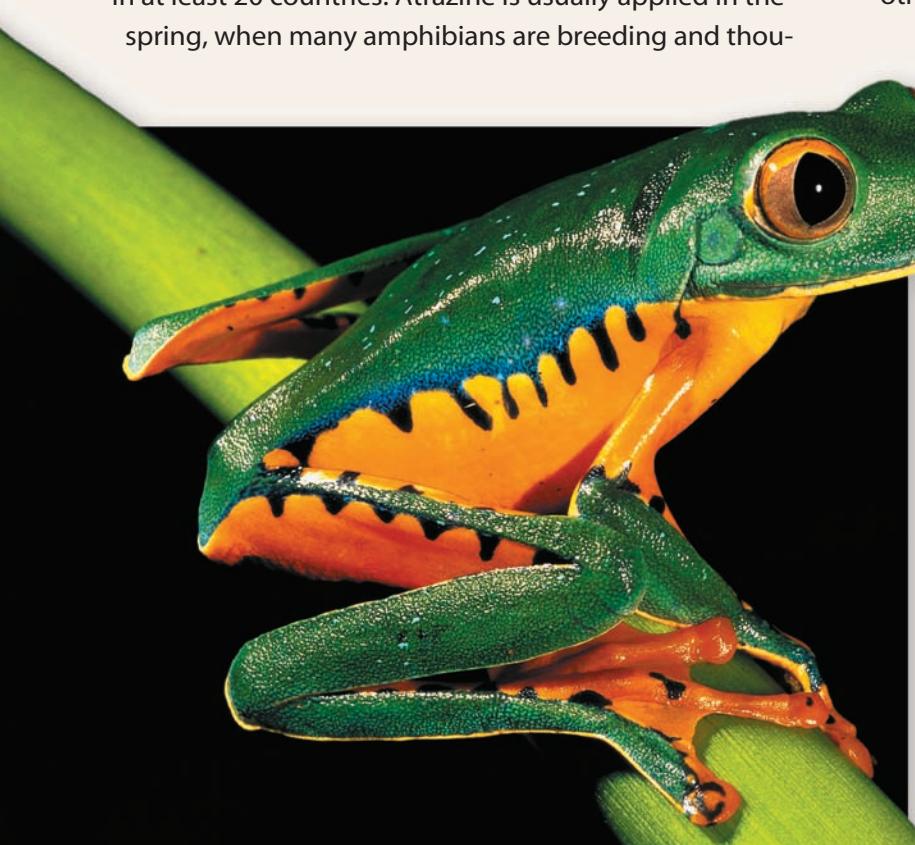
Hayes focused on atrazine, the most widely used herbicide in the world and a common contaminant in fresh water. More than 70 million pounds of atrazine are applied to farmland in the United States every year, and it is used in at least 20 countries. Atrazine is usually applied in the spring, when many amphibians are breeding and thou-

sands of tadpoles swim in the ditches, ponds, and streams that receive runoff from farms.

In his laboratory, Hayes and his associates raised frog tadpoles in water containing no atrazine and in water with concentrations ranging from 0.01 parts per billion (ppb) up to 25 ppb. The U.S. Environmental Protection Agency considers environmental levels of atrazine of 10 to 20 ppb of no concern; the level it considers safe in drinking water is 3 ppb. Rainwater in Iowa has been measured to contain 40 ppb. In Switzerland, where the use of atrazine is illegal, the chemical has been measured at approximately 1 ppb in rainwater.

In the Hayes laboratory, concentrations as low as 0.1 ppb had a dramatic effect on tadpole development: it feminized the males. In some of the adult males that developed from these larvae, the vocal structures used in mating calls were smaller than normal, female sex organs developed, and eggs were found growing in the testes. In other studies, normal adult male frogs exposed to 25 ppb had a tenfold reduction in testosterone levels and did not produce sperm. You can imagine the disastrous effects these developmental and hormonal changes could have on the capacity of frogs to breed and reproduce.

But Hayes's experiments were performed in the laboratory, with a species of frog bred for laboratory use. Would his results be the same in nature? To find out, he and his students traveled from Utah to Iowa, sampling water and collecting frogs. They analyzed the water



Frogs Are Having Serious Problems An alarming number of species of frogs, such as this tiny leaf frog (*Agalychnis calcarifer*) from Ecuador, are in danger of becoming extinct. The numerous possible reasons for the decline in global amphibian populations have been a subject of widespread scientific investigation.



A Biologist at Work Tyrone Hayes grew up near the great Congaree Swamp in South Carolina collecting turtles, snakes, frogs, and toads. Now a professor of biology at the University of California at Berkeley, he has more than 3,000 frogs in his laboratory and studies hormonal control of their development.

for atrazine and examined the frogs. In the only site where atrazine was undetectable in the water, the frogs were normal; in all the other sites, male frogs had abnormalities of the sex organs.

Like other biologists, Hayes made observations. He then made predictions based on those observations, and designed and carried out experiments to test his predictions. Some of the conclusions from his experiments, described at the end of this chapter, could have profound implications not only for amphibians but also for other animals, including humans.

IN THIS CHAPTER we identify and examine the most common features of living organisms and put those features into the context of the major principles that underlie all biology. Next we offer a brief outline of how life evolved and how the different organisms on Earth are related. We then turn to the subjects of biological inquiry and the scientific method. Finally we consider how knowledge discovered by biologists influences public policy.

CHAPTER OUTLINE

- 1.1 What Is Biology?
- 1.2 How Is All Life on Earth Related?
- 1.3 How Do Biologists Investigate Life?
- 1.4 How Does Biology Influence Public Policy?

1.1 What Is Biology?

Biology is the scientific study of living things. Biologists define “living things” as all the diverse organisms descended from a single-celled ancestor that evolved almost 4 billion years ago. Because of their common ancestry, living organisms share many characteristics that are not found in the nonliving world. Living organisms:

- consist of one or more cells
- contain genetic information
- use genetic information to reproduce themselves
- are genetically related and have evolved
- can convert molecules obtained from their environment into new biological molecules
- can extract energy from the environment and use it to do biological work
- can regulate their internal environment

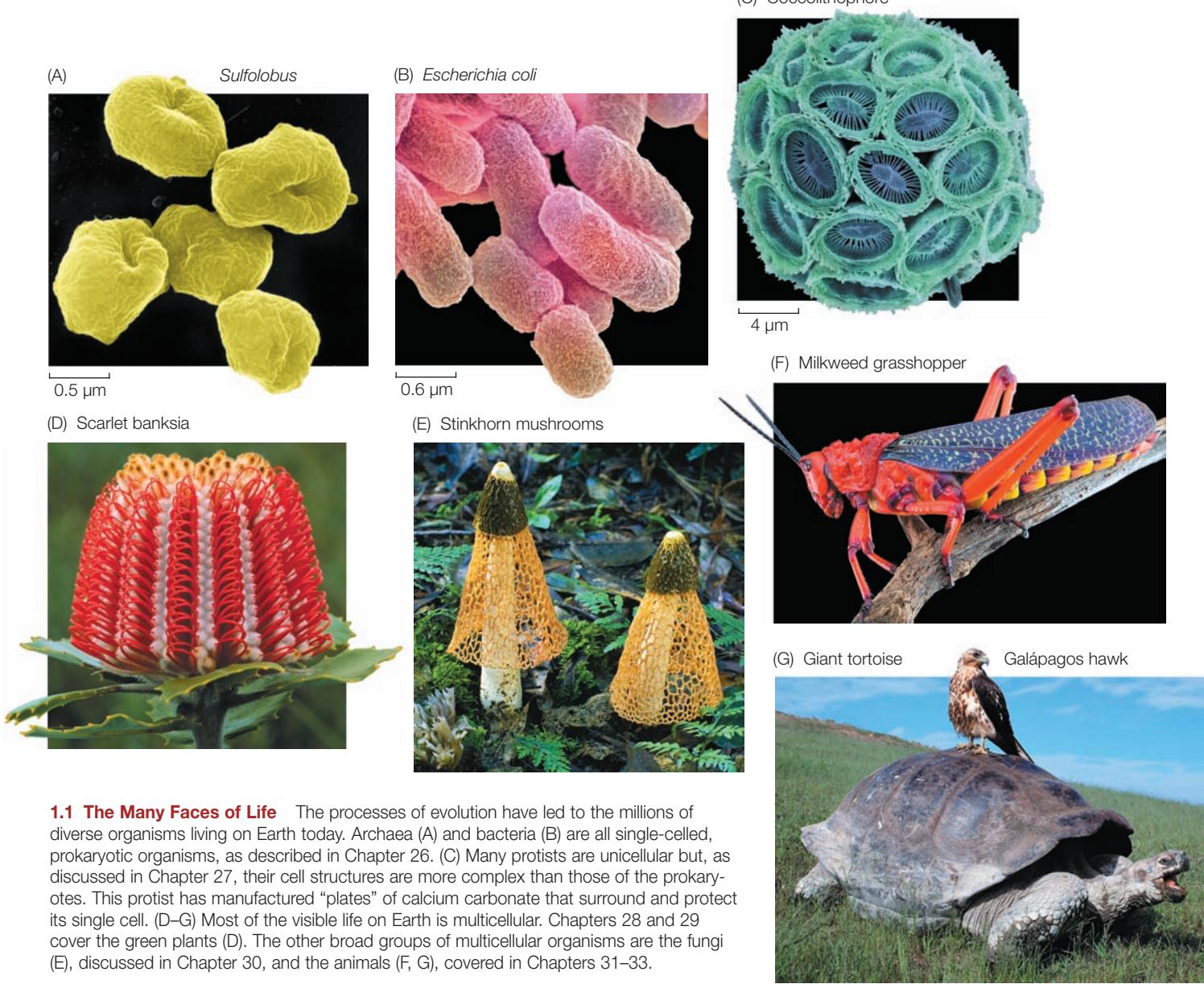
This simple list, however, belies the incredible complexity and diversity of life. Some forms of life may not display all of these characteristics all of the time. For example, the seed of a desert plant may go for many years without extracting energy from the environment, converting molecules, regulating its internal environment, or reproducing; yet the seed is alive.

And what about viruses? Viruses do not consist of cells, and they cannot carry out physiological functions on their own; they must parasitize host cells to do those jobs for them. Yet viruses contain genetic information, and they certainly mutate and evolve (as we know, because evolving flu viruses require constant changes in the vaccines we create to combat them). The existence of viruses depends on cells, and it is highly probable that viruses evolved from cellular life forms. So, are viruses alive? What do you think?

This book explores the characteristics of life, how these characteristics vary among organisms, how they evolved, and how they work together to enable organisms to survive and reproduce. *Evolution* is a central theme of biology and therefore of this book. Through differential survival and reproduction, living systems evolve and become adapted to Earth’s many environments. The processes of evolution have generated the enormous diversity that we see today as life on Earth.

Cells are the basic unit of life

We lay the chemical foundation for our study of life in the next three chapters, after which we will turn to cells and the processes by which they live, reproduce, age, and die. Some organisms are *unicellular*, consisting of a single cell that carries out



1.1 The Many Faces of Life The processes of evolution have led to the millions of diverse organisms living on Earth today. Archaea (A) and bacteria (B) are all single-celled, prokaryotic organisms, as described in Chapter 26. (C) Many protists are unicellular but, as discussed in Chapter 27, their cell structures are more complex than those of the prokaryotes. This protist has manufactured “plates” of calcium carbonate that surround and protect its single cell. (D–G) Most of the visible life on Earth is multicellular. Chapters 28 and 29 cover the green plants (D). The other broad groups of multicellular organisms are the fungi (E), discussed in Chapter 30, and the animals (F, G), covered in Chapters 31–33.

all the functions of life (**Figure 1.1A–C**). Others are *multicellular*, made up of many cells that are specialized for different functions (**Figure 1.1D–G**). Viruses are *acellular*, although they depend on cellular organisms.

The discovery of cells was made possible by the invention of the microscope in the 1590s by the Dutch spectacle makers Hans and Zaccharias Janssen (father and son). In the mid- to late 1600s, Antony van Leeuwenhoek of Holland and Robert Hooke of England both made improvements on the Janssens' technology and used it to study living organisms. Van Leeuwenhoek discovered that drops of pond water teemed with single-celled organisms, and he made many other discoveries as he progressively improved his microscopes over a long lifetime of research. Hooke put pieces of plants under his microscope and observed that they were made up of repeated units he called *cells* (**Figure 1.2**). In 1676, Hooke wrote that van Leeuwenhoek had observed “a vast number of small animals in his Excrements which were most abounding when he was troubled with a Loosenesse and very few or none when he was well.” This simple observation

represents the discovery of bacteria—and makes one wonder why scientists do some of the things they do.

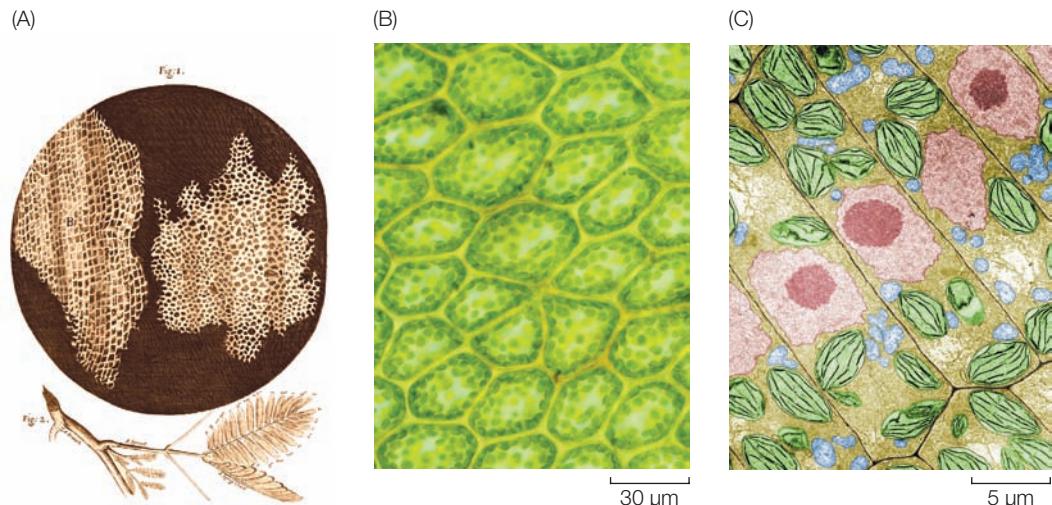
More than a hundred years passed before studies of cells advanced significantly. As they were dining together one evening in 1838, Matthias Schleiden, a German biologist, and Theodor Schwann, from Belgium, discussed their work on plant and animal tissues, respectively. They were struck by the similarities in their observations and came to the conclusion that the basic structural elements of plants and animals were essentially the same. They formulated their conclusion as the **cell theory**, which states that:

- Cells are the basic structural and physiological units of all living organisms.
- Cells are both distinct entities and building blocks of more complex organisms.

But Schleiden and Schwann also believed (wrongly) that cells emerged by the self-assembly of nonliving materials, much as crystals form in a solution of salt. This conclusion was in ac-

1.2 Cells Are the Building Blocks of Life

The development of microscopes revealed the microbial world to seventeenth-century scientists such as Robert Hooke, who proposed the concept of cells based on his observations. (A) Hooke drew the cells of a slice of plant tissue (cork) as he saw them under his optical microscope. (B) A modern optical, or “light,” microscope reveals the intricacies of cells in a leaf. (C) Transmission electron microscopes (TEMs) allow scientists to see even smaller objects. TEMs do not visualize color; here color has been added to a black-and-white micrograph of cells in a duckweed stem.



cordance with the prevailing view of the day, which was that life can arise from non-life by spontaneous generation—mice from dirty clothes, maggots from dead meat, or insects from pond water.

The debate continued until 1859, when the French Academy of Sciences sponsored a contest for the best experiment to prove or disprove spontaneous generation. The prize was won by the great French scientist Louis Pasteur, who demonstrated that sterile broth directly exposed to the dirt and dust in air developed a culture of microorganisms, but a similar container of broth not directly exposed to air remained sterile (see Figure 4.7). Pasteur’s experiment did not prove that it was microorganisms in the air that caused the broth to become infected, but it did uphold the conclusion that life must be present in order for new life to be generated.

Today scientists accept the fact that all cells come from pre-existing cells and that the functional properties of organisms derive from the properties of their cells. Since cells of all kinds share both essential mechanisms and a common ancestry that goes back billions of years, modern cell theory has additional elements:

- All cells come from preexisting cells.
- All cells are similar in chemical composition.
- Most of the chemical reactions of life occur in aqueous solution within cells.
- Complete sets of genetic information are replicated and passed on during cell division.
- Viruses lack cellular structure but remain dependent on cellular organisms.

At the same time Schleiden and Schwann were building the foundation for the cell theory, Charles Darwin was beginning to understand how organisms undergo evolutionary change.

All of life shares a common evolutionary history

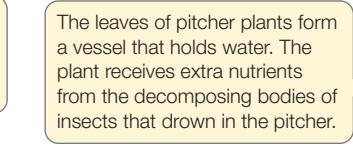
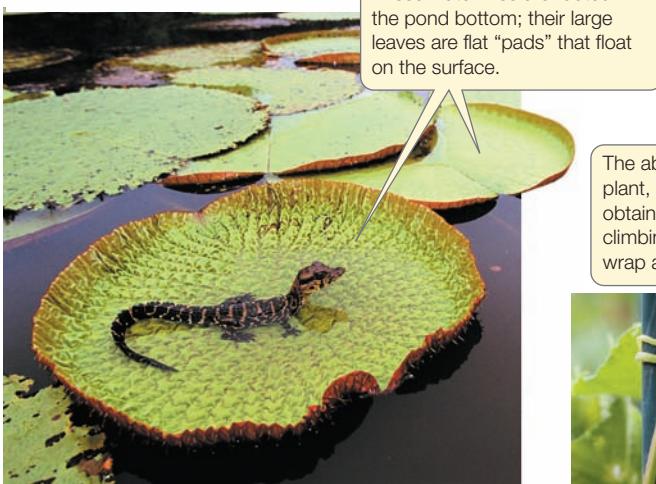
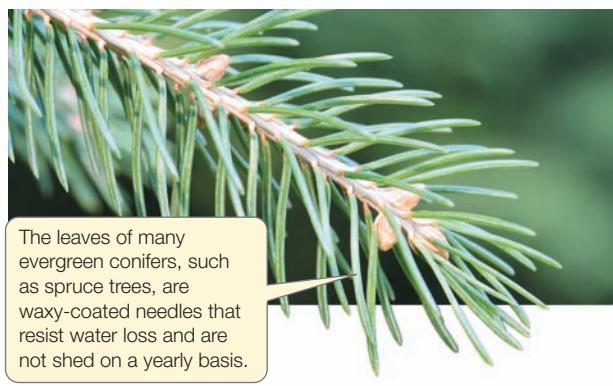
Evolution—change in the genetic makeup of biological populations through time—is the major unifying principle of biology.

Charles Darwin compiled factual evidence for evolution in his 1859 book *On the Origin of Species*. Since then, biologists have gathered massive amounts of data supporting Darwin’s theory that all living organisms are descended from a common ancestor. Darwin also proposed one of the most important processes that produce evolutionary change. He argued that differential survival and reproduction among individuals in a population, which he termed **natural selection**, could account for much of the evolution of life.

Although Darwin proposed that living organisms are descended from common ancestors and are therefore related to one another, he did not have the advantage of understanding the mechanisms of genetic inheritance. Even so, he observed that offspring resembled their parents; therefore, he surmised, such mechanisms had to exist. That simple fact is the basis for the concept of a **species**. Although the precise definition of a species is complicated, in its most widespread usage it refers to a group of organisms that can produce viable and fertile offspring with one another.

But offspring do differ from their parents. Any population of a plant or animal species displays variation, and if you select breeding pairs on the basis of some particular trait, that trait is more likely to be present in their offspring than in the general population. Darwin himself bred pigeons, and was well aware of how pigeon fanciers selected breeding pairs to produce offspring with unusual feather patterns, beak shapes, or body sizes (see Figure 21.2). He realized that if humans could select for specific traits in domesticated animals, the same process could operate in nature; hence the term *natural selection* as opposed to artificial (human-imposed) selection.

How would natural selection function? Darwin postulated that different probabilities of survival and reproductive success would do the job. He reasoned that the reproductive capacity of plants and animals, if unchecked, would result in unlimited growth of populations, but we do not observe such growth in nature; in most species, only a small percentage of offspring survive to reproduce. Thus any trait that confers even a small increase in the probability that its possessor will survive and reproduce would be spread in the population.



1.3 Adaptations to the Environment The leaves of all plants are specialized for photosynthesis—the sun-light-powered transformation of water and carbon dioxide into larger structural molecules called carbohydrates. The leaves of different plants, however, display many different adaptations to their individual environments.

Because organisms with certain traits survive and reproduce best under specific sets of conditions, natural selection leads to **adaptations**: structural, physiological, or behavioral traits that enhance an organism's chances of survival and reproduction in its environment (Figure 1.3). In addition to natural selection, evolutionary processes such as sexual selection (selection due to mate choice) and genetic drift (the random fluctuation of gene frequencies in a population due to chance events) contribute to the rise of diverse adaptations. These processes operating over evolutionary history have led to the remarkable array of life on Earth.

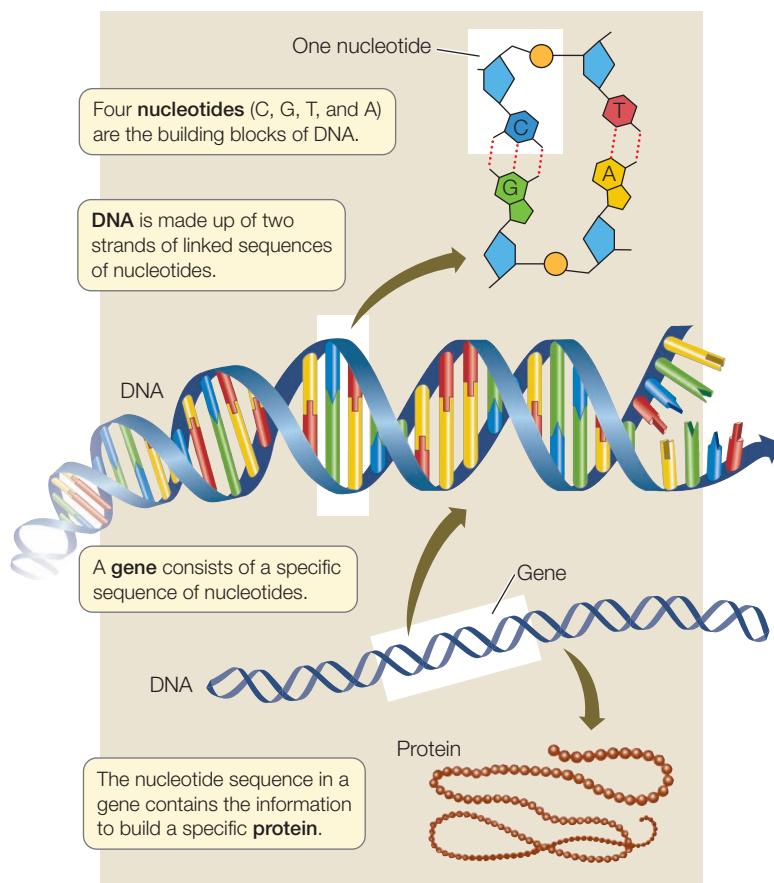
If all cells come from preexisting cells, and if all the diverse species of organisms on Earth are related by descent with modification from a common ancestor, then what is the source of information that is passed from parent to daughter cells and from parental organisms to their offspring?

Biological information is contained in a genetic language common to all organisms

Cells are the basic building blocks of organisms, but even a single cell is complex, with many internal structures and many functions that depend on information. The information required

for a cell to function and interact with other cells—the “blueprint” for existence—is contained in the cell’s **genome**, the sum total of all the DNA molecules it contains. **DNA** (deoxyribonucleic acid) molecules are long sequences of four different sub-units called **nucleotides**. The sequence of the nucleotides contains genetic information. **Genes** are specific segments of DNA encoding the information the cell uses to make **proteins** (Figure 1.4). Protein molecules govern the chemical reactions within cells and form much of an organism’s structure.

By analogy with a book, the nucleotides of DNA are like the letters of an alphabet. Protein molecules are the sentences. Combinations of proteins that form structures and control biochemical processes are the paragraphs. The structures and processes that are organized into different systems with specific tasks (such as digestion or transport) are the chapters of the book, and the complete book is the organism. If you were to write out your own genome using four letters to represent the four nucleotides, you would write more than 3 billion letters. Using the size type you are reading now, your genome would fill about a thousand books the size of this one. The mechanisms of evolution, including natural selection, are the authors and editors of all the books in the library of life.



1.4 DNA Is Life's Blueprint The instructions for life are contained in the sequences of nucleotides in DNA molecules. Specific DNA nucleotide sequences comprise genes. The average length of a single human gene is 16,000 nucleotides. The information in each gene provides the cell with the information it needs to manufacture molecules of a specific protein.

All the cells of a multicellular organism contain the same genome, yet different cells have different functions and form different structures—contractile proteins form in muscle cells, hemoglobin in red blood cells, digestive enzymes in gut cells, and so on. Therefore, different types of cells in an organism must express different parts of the genome. How cells control gene expression in ways that enable a complex organism to develop and function is a major focus of current biological research.

The genome of an organism consists of thousands of genes. If the nucleotide sequence of a gene is altered, it is likely that the protein that gene encodes will be altered. Alterations of the genome are called *mutations*. Mutations occur spontaneously; they can also be induced by outside factors, including chemicals and radiation. Most mutations are either harmful or have no effect, but occasionally a mutation improves the functioning of the organism under the environmental conditions it encounters. Such beneficial mutations are the raw material of evolution and lead to adaptations.

Cells use nutrients to supply energy and to build new structures

Living organisms acquire *nutrients* from the environment. Nutrients supply the organism with energy and raw materials for carrying out biochemical reactions. Life depends on thousands

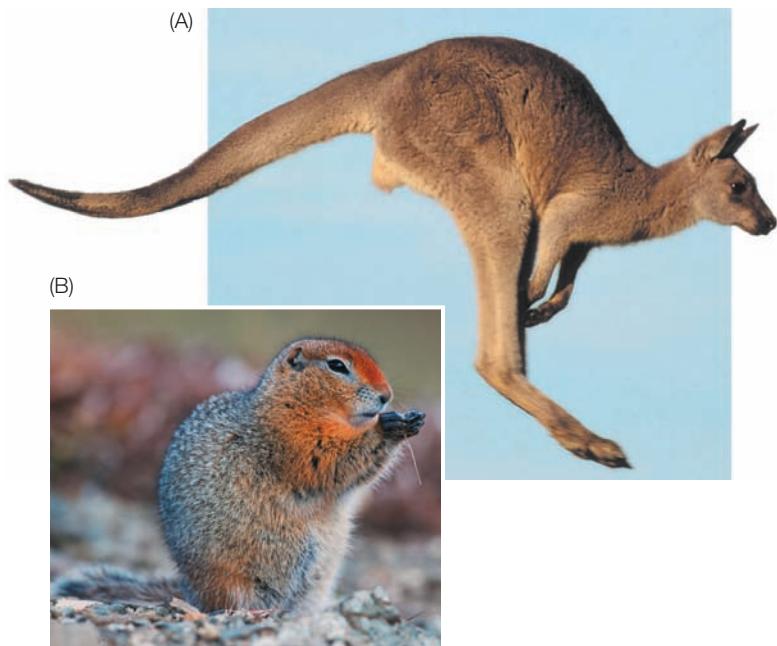
of biochemical reactions that occur inside cells. Some of these reactions break down nutrient molecules into smaller chemical units, and in the process some of the energy contained in the chemical bonds of the nutrients is captured by high-energy molecules that can be used to do different kinds of cellular work.

One obvious kind of work cells do is mechanical—moving molecules from one cellular location to another, moving whole cells or tissues, or even moving the organism itself, as muscles do (Figure 1.5A). The most basic cellular work is the building, or *synthesis*, of new complex molecules and structures from smaller chemical units. For example, we are all familiar with the fact that carbohydrates eaten today may be deposited in the body as fat tomorrow (Figure 1.5B). Still another kind of work is the electrical work that is the essence of information processing in nervous systems. The sum total of all the chemical transformations and other work done in all the cells of an organism is its **metabolism**, or **metabolic rate**.

The myriad of biochemical reactions that go on in cells are integrally linked in that the products of one are the raw materials of the next. These complex networks of reactions must be integrated and precisely controlled; when they are not, the result is disease.

Living organisms regulate their internal environment

Multicellular organisms have an *internal environment* that is not cellular. That is, their individual cells are bathed in extracellular fluids, from which they receive nutrients and into which they excrete waste products of metabolism. The cells of multicellu-



1.5 Energy Can Be Used Immediately or Stored (A) Animal cells break down and release the energy contained in the chemical bonds of food molecules to do mechanical work—in this kangaroo's case, to jump. (B) The cells of this Arctic ground squirrel have broken down the complex carbohydrates in plants and converted their molecules into fats, which are stored in the animal's body to provide an energy supply for the cold months.

lar organisms are specialized, or *differentiated*, to contribute in some way to the maintenance of the internal environment. With the evolution of specialization, differentiated cells lost many of the functions carried out by single-celled organisms, and must depend on the internal environment for essential services.

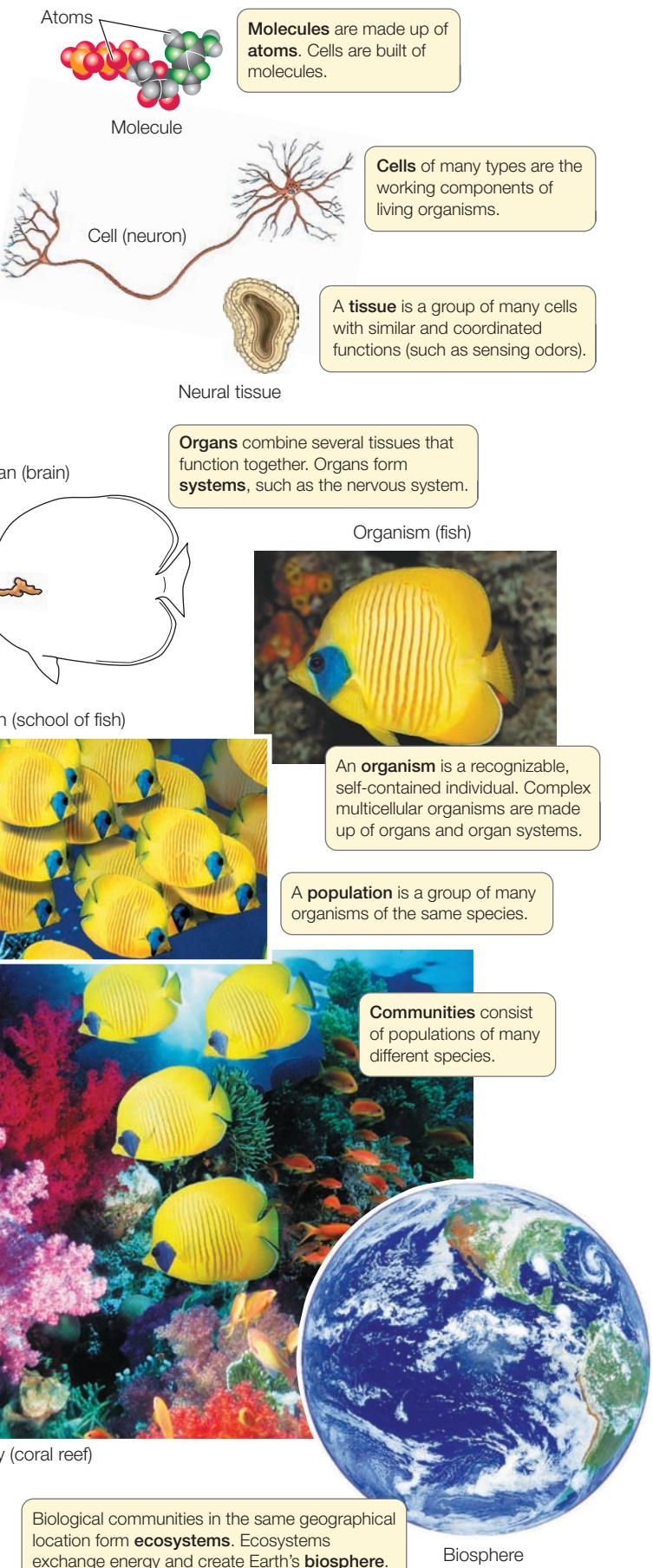
To accomplish their specialized tasks, assemblages of differentiated cells are organized into *tissues*. For example, a single muscle cell cannot generate much force, but when many cells combine to form the tissue of a working muscle, considerable force and movement can be generated (see Figure 1.5B). Different tissue types are organized to form *organs* that accomplish specific functions. For example, the heart, brain, and stomach are each constructed of several types of tissues. Organs whose functions are interrelated can be grouped into *organ systems*; the stomach, intestine, and esophagus, for example, are parts of the digestive system. The functions of cells, tissues, organs, and organ systems are all integral to the multicellular *organism*. We cover the biology of organisms in Parts Eight and Nine of this book.

Living organisms interact with one another

The internal hierarchy of the individual organism is matched by the external hierarchy of the biological world (Figure 1.6). Organisms do not live in isolation. A group of individuals of the same species that interact with one another is a *population*, and populations of all the species that live and interact in the same area are called a *community*. Communities together with their abiotic environment constitute an *ecosystem*.

Individuals in a population interact in many different ways. Animals eat plants and other animals (usually members of another species) and compete with other species for food and other resources. Some animals will prevent other individuals of their own species from exploiting a resource, whether it be food, nesting sites, or mates. Animals may also *cooperate* with members of their species, forming social units such as a termite colony or a flock of birds. Such interactions have resulted in the evolution of social behaviors such as communication.

Plants also interact with their external environment, which includes other plants, animals, and microorganisms. All terrestrial plants depend on complex partnerships with fungi, bacteria, and animals. Some of these partnerships are necessary to obtain nutrients, some to produce fertile seeds, and still others to disperse seeds. Plants compete with each other



1.6 Biology Is Studied at Many Levels of Organization

Life's properties emerge when DNA and other molecules are organized in cells. Energy flows through all the biological levels shown here.

for light and water, and they have ongoing evolutionary interactions with the animals that eat them, evolving anti-predation adaptations or ways to attract the animals that assist in their reproduction. The interactions of populations of different plant and animal species in a community are major evolutionary forces that produce specialized adaptations.

Communities interacting over a broad geographic area with distinguishing physical features form ecosystems; examples might include an Arctic tundra, a coral reef, or a tropical rainforest. The ways in which species interact with one another and with their environment in communities and in ecosystems is the subject of *ecology* and of Part Ten of this book.

Discoveries in biology can be generalized

Because all life is related by descent from a common ancestor, shares a genetic code, and consists of similar building blocks—cells—knowledge gained from investigations of one type of organism can, with care, be generalized to other organisms. Biologists use **model systems** for research, knowing that they can extend their findings to other organisms, including humans. For example, our basic understanding of the chemical reactions in cells came from research on bacteria but is applicable to all cells, including those of humans. Similarly, the biochemistry of photosynthesis—the process by which plants use sunlight to produce biological molecules—was largely worked out from experiments on *Chlorella*, a unicellular green alga (see Figure 10.13). Much of what we know about the genes that control plant development is the result of work on *Arabidopsis thaliana*, a relative of the mustard plant. Knowledge about how animals develop has come from work on sea urchins, frogs, chickens, roundworms, and fruit flies. And recently, the discovery of a major gene controlling human skin color came from work on zebrafish. Being able to generalize from model systems is a powerful tool in biology.

1.1 RECAP

Living organisms are made of (or depend on) cells, are related by common descent and evolve, contain genetic information and use it to reproduce, extract energy from their environment and use it to do biological work, synthesize complex molecules to construct biological structures, regulate their internal environment, and interact with one another.

- Describe the relationship between evolution by natural selection and the genetic code. **See pp. 6–7**
- Why can the results of biological research on one species often be generalized to very different species? **See p. 9**

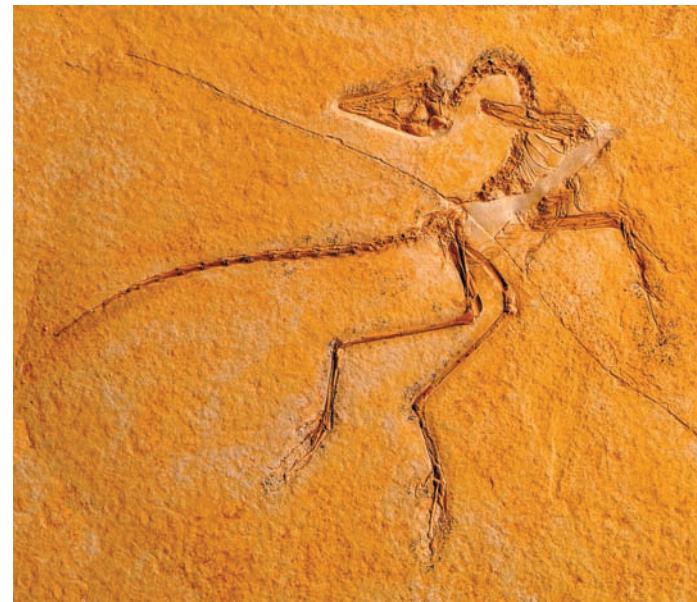
Now that you have an overview of the major features of life that you will explore in depth in this book, you can ask how and when life first emerged. In the next section we will summarize briefly the history of life from the earliest simple life forms to the complex and diverse organisms that inhabit our planet today.

1.2 How Is All Life on Earth Related?

What do biologists mean when they say that all organisms are *genetically related*? They mean that species on Earth share a *common ancestor*. If two species are similar, as dogs and wolves are, then they probably have a common ancestor in the fairly recent past. The common ancestor of two species that are more different—say, a dog and a deer—probably lived in the more distant past. And if two organisms are very different—such as a dog and a clam—then we must go back to the *very* distant past to find their common ancestor. How can we tell how far back in time the common ancestor of any two organisms lived? In other words, how do we discover the evolutionary relationships among organisms?

For many years, biologists have investigated the history of life by studying the *fossil record*—the preserved remains of organisms that lived in the distant past (Figure 1.7). Geologists supplied knowledge about the ages of fossils and the nature of the environments in which they lived. Biologists then inferred the evolutionary relationships among living and fossil organisms by comparing their anatomical similarities and differences. Frequently big gaps existed in the fossil record, forcing biologists to predict the nature of the “missing links” between two lineages of organisms. As the fossil record became more complete, those missing links were filled in.

Molecular methods for comparing genomes, described in Chapter 24, are enabling biologists to more accurately establish the degrees of relationship between living organisms and to use that information to interpret the fossil record. Molecular information can occasionally be gleaned from fossil specimens, such as recently deciphered genetic material from fossil bones of Ne-



1.7 Fossils Give Us a View of Past Life This fossil, formed some 150 million years ago, is that of an *Archaeopteryx*, the earliest known representative of the birds. Birds evolved from the same group of reptiles as the modern crocodiles.

anderthals that led to the conclusion that even though Neanderthals and modern humans coexisted, they did not interbreed.

In general, the greater the differences between the genomes of two species, the more distant their common ancestor. Using molecular techniques, biologists are exploring fundamental questions about life. What were the earliest forms of life? How did simple organisms give rise to the great diversity of organisms alive today? Can we reconstruct a family tree of life?

Life arose from non-life via chemical evolution

Geologists estimate that Earth formed between 4.6 and 4.5 billion years ago. At first, the planet was not a very hospitable place. It was some 600 million years or more before the earliest life evolved. If we picture the history of Earth as a 30-day month, life first appeared somewhere toward the end of the first week (**Figure 1.8**).

When we consider how life might have arisen from nonliving matter, we must take into account the properties of the young

Earth's atmosphere, oceans, and climate, all of which were very different than they are today. Biologists postulate that complex biological molecules first arose through the random physical association of chemicals in that environment. Experiments simulating the conditions on early Earth have confirmed that the generation of complex molecules under such conditions is possible, even probable. The critical step for the evolution of life, however, had to be the appearance of molecules that could reproduce themselves and also serve as templates for the synthesis of large molecules with complex but stable shapes. The variation of the shapes of these large, stable molecules (described in Chapters 3 and 4) enabled them to participate in increasing numbers and kinds of chemical reactions with other molecules.

Cellular structure evolved in the common ancestor of life

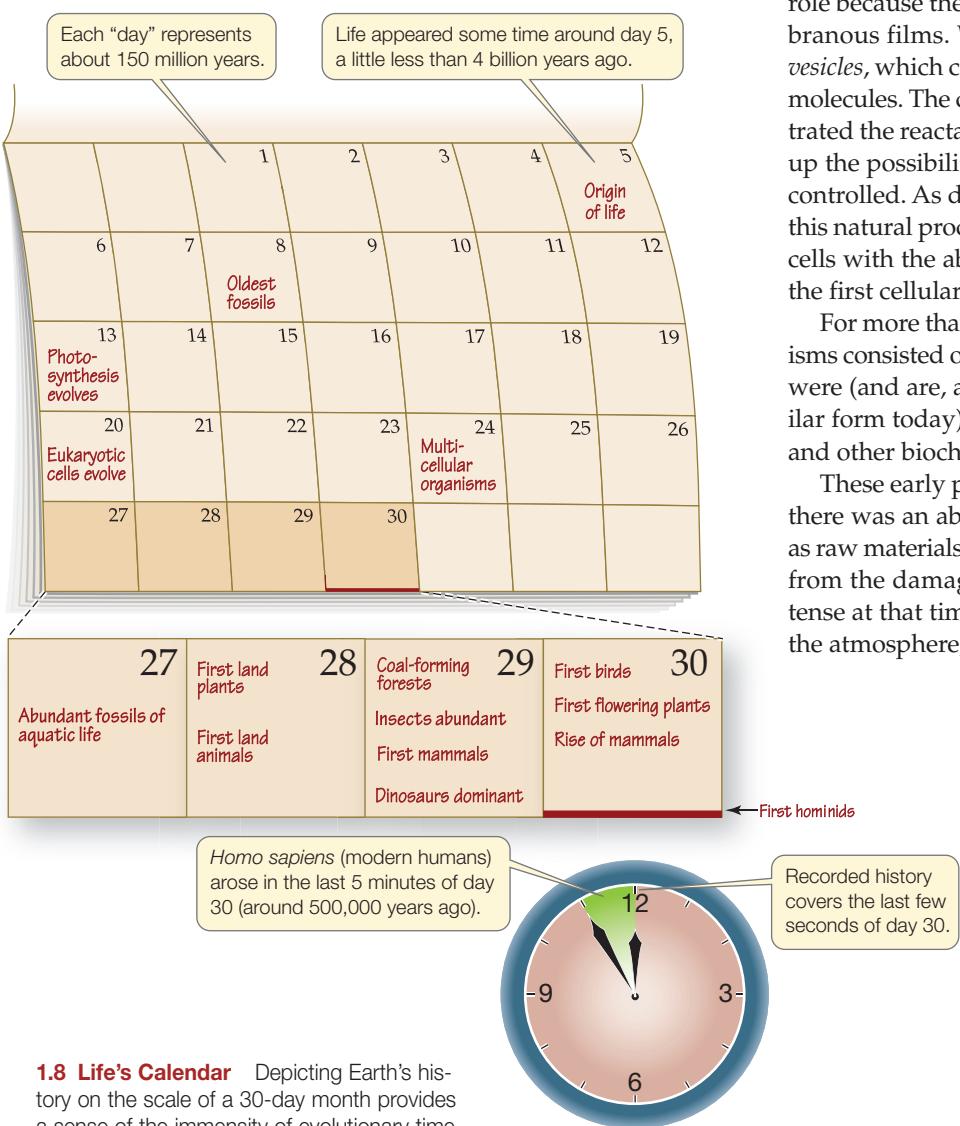
The second critical step in the origin of life was the enclosure of complex biological molecules by *membranes* that contained them in a compact internal environment separate from the surrounding external environment. Fatlike molecules played a critical role because they are not soluble in water and they form membranous films. When agitated, these films can form spherical *vesicles*, which could have enveloped assemblages of biological molecules. The creation of an internal environment that concentrated the reactants and products of chemical reactions opened up the possibility that those reactions could be integrated and controlled. As described in Section 4.4, scientists postulate that this natural process of membrane formation resulted in the first cells with the ability to replicate themselves—the evolution of the first cellular organisms.

For more than 2 billion years after cells originated, all organisms consisted of only one cell. These first unicellular organisms were (and are, as multitudes of their descendants exist in similar form today) **prokaryotes**. Prokaryotic cells consist of DNA and other biochemicals enclosed in a membrane.

These early prokaryotes were confined to the oceans, where there was an abundance of complex molecules they could use as raw materials and sources of energy. The ocean shielded them from the damaging effects of ultraviolet light, which was intense at that time because there was little or no oxygen (O_2) in the atmosphere, and hence no protective ozone (O_3) layer.

Photosynthesis changed the course of evolution

To fuel their cellular metabolism, the earliest prokaryotes took in molecules directly from their environment and broke these small molecules down to release and use the energy contained in their chemical bonds. Many modern species of prokaryotes still function this way, and very successfully. During the early eons of life on Earth, there was no oxygen in the atmosphere. In fact, oxygen was toxic to the life forms that existed then.





1.9 Photosynthetic Organisms Changed Earth's Atmosphere

These strands are composed of many cells of cyanobacteria. This modern species (*Oscillatoria tenuis*) may be very similar to the early photosynthetic prokaryotes responsible for the buildup of oxygen in Earth's atmosphere.

About 2.7 billion years ago, the evolution of **photosynthesis** changed the nature of life on Earth. The chemical reactions of photosynthesis transform the energy of sunlight into a form of biological energy that can power the synthesis of large molecules (see Chapter 10). These large molecules are the building blocks of cells, and they can be broken down to provide metabolic energy. Photosynthesis is the basis of much of life on Earth today because its energy-capturing processes provide food for other organisms.

Early photosynthetic cells were probably similar to present-day prokaryotes called *cyanobacteria* (Figure 1.9). Over time, photosynthetic prokaryotes became so abundant that vast quantities of O₂, which is a by-product of photosynthesis, slowly began to accumulate in the atmosphere. Oxygen was poisonous to many of the prokaryotes that lived at that time. Those organisms that did tolerate oxygen, however, were able to proliferate as the presence of oxygen opened up vast new avenues of evolution. *Aerobic metabolism* (energy production based on the conversion of O₂) is more efficient than *anaerobic* (non-O₂-using) *metabolism*, and today it is used by the majority of Earth's organisms. Aerobic metabolism allowed cells to grow larger.

Oxygen in the atmosphere also made it possible for life to move onto land. For most of life's history, ultraviolet (UV) radiation falling on Earth's surface was too intense to allow life to exist outside the shielding water. But the accumulation of photosynthetically generated oxygen in the atmosphere for more than 2 billion years gradually produced a layer of ozone in the upper atmosphere. By about 500 million years ago, the ozone layer was sufficiently dense and absorbed enough UV radiation to make it possible for organisms to leave the protection of the water and live on land.

Eukaryotic cells evolved from prokaryotes

Another important step in the history of life was the evolution of cells with discrete intracellular compartments, called **organelles**, which were capable of taking on specialized cellular functions. This event happened about 3 weeks into our calendar of Earth's history (see Figure 1.8). One of these organelles, the dense-appearing *nucleus* (Latin *nux*, "nut" or "core"), came to contain the cell's genetic information and gives these cells their name: **eukaryotes** (Greek *eu*, "true"; *karyon*, "kernel" or "core"). The eukaryotic cell is completely distinct from the cells of prokaryotes (*pro*, "before"), which lack nuclei and other internal compartments.

Some organelles are hypothesized to have originated by **endosymbiosis** when cells ingested smaller cells. The *mitochondria* that generate a cell's energy probably evolved from engulfed prokaryotic organisms. And *chloroplasts*—organelles specialized to conduct photosynthesis—could have originated when photosynthetic prokaryotes were ingested by larger eukaryotes. If the larger cell failed to break down this intended food object, a partnership could have evolved in which the ingested prokaryote provided the products of photosynthesis and the host cell provided a good environment for its smaller partner.

Multicellularity arose and cells became specialized

Until just over a billion years ago, all the organisms that existed—whether prokaryotic or eukaryotic—were unicellular. An important evolutionary step occurred when some eukaryotes failed to separate after cell division, remaining attached to each other. The permanent association of cells made it possible for some cells to specialize in certain functions, such as reproduction, while other cells specialized in other functions, such as absorbing nutrients and distributing them to neighboring cells. This **cellular specialization** enabled multicellular eukaryotes to increase in size and become more efficient at gathering resources and adapting to specific environments.

Biologists can trace the evolutionary tree of life

If all the species of organisms on Earth today are the descendants of a single kind of unicellular organism that lived almost 4 billion years ago, how have they become so different? A simplified answer is that as long as individuals within a population mate with one another, structural and functional changes can evolve within that population, but the population will remain one species. However, if something happens to isolate some members of a population from the others, the structural and functional differences between the two groups may accumulate over time. The two groups may diverge to the point where their members can no longer reproduce with each other and are thus distinct species. We discuss this evolutionary process, called *speciation*, in Chapter 23.

Biologists give each species a distinctive scientific name formed from two Latinized names (a **binomial**). The first name identifies the species' *genus*—a group of species that share a recent common ancestor. The second is the name of the species. For

example, the scientific name for the human species is *Homo sapiens*: *Homo* is our genus and *sapiens* our species. *Homo* is Latin for “man”; *sapiens* is from the Latin for word for “wise” or “rational.”

Tens of millions of species exist on Earth today. Many times that number lived in the past but are now extinct. Many millions of speciation events created this vast diversity, and the unfolding of these events can be diagrammed as an evolutionary “tree” whose branches describe the order in which populations split and eventually evolved into new species, as described in Chapter 22. Much of biology is based on comparisons among species, and these comparisons are useful precisely because we can place species in an evolutionary context relative to one another. Our ability to do this has been greatly enhanced in recent decades by our ability to sequence and compare the genomes of different species.

Genome sequencing and other molecular techniques have allowed *systematists*—scientists who study the evolution and classification of life’s diverse organisms—to augment evolutionary knowledge based on the fossil record with a vast array of molecular evidence. The result is the ongoing compilation of *phylogenetic trees* that document and diagram evolutionary relationships as part of an overarching tree of life, the broadest categories of which are shown in **Figure 1.10**. (The tree is expanded in this book’s Appendix; you can also explore the tree interactively at <http://tolweb.org/tree/>.)

Although many details remain to be clarified, the broad outlines of the tree of life have been determined. Its branching patterns are based on a rich array of evidence from fossils, struc-

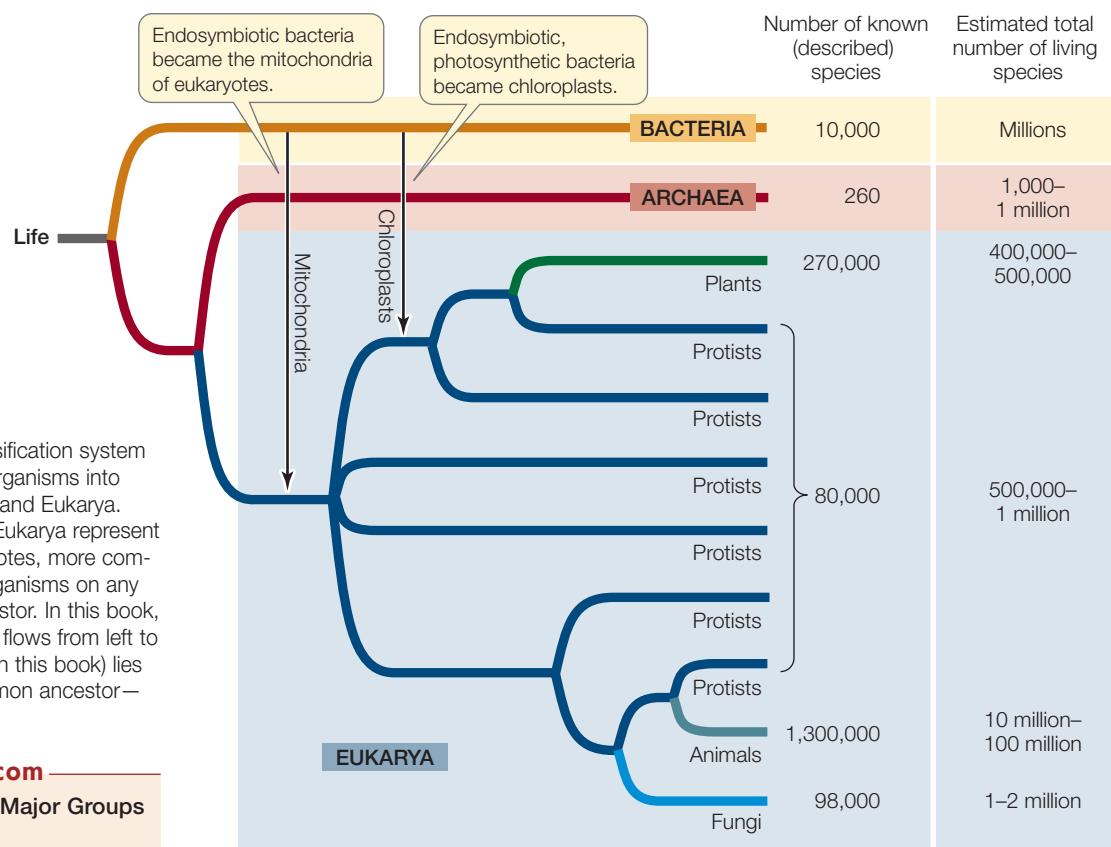
tures, metabolic processes, behavior, and molecular analyses of genomes. Molecular data in particular have been used to separate the tree into three major **domains**: Archaea, Bacteria, and Eukarya. The organisms of each domain have been evolving separately from those in the other domains for more than a billion years.

Organisms in the domains **Archaea** and **Bacteria** are single-celled prokaryotes. However, members of these two groups differ so fundamentally in their metabolic processes that they are believed to have separated into distinct evolutionary lineages very early. Species belonging to the third domain—**Eukarya**—have eukaryotic cells whose mitochondria and chloroplasts may have originated from the ingestion of prokaryotic cells, as described on page 11.

The three major groups of multicellular eukaryotes—plants, fungi, and animals—each evolved from a different group of the eukaryotes generally referred to as *protists*. The chloroplast-containing, photosynthetic protist that gave rise to plants was completely distinct from the protist that was ancestral to both animals and fungi, as can be seen from the branching pattern of Figure 1.10. Although most protists are unicellular (and thus sometimes called *microbial eukaryotes*), multicellularity has evolved in several protist lineages.

The tree of life is predictive

There are far more species alive on Earth than biologists have discovered and described to date. In fact, most species on Earth



1.10 The Tree of Life The classification system used in this book divides Earth’s organisms into three domains: Bacteria, Archaea, and Eukarya. The darkest blue branches within Eukarya represent various groups of microbial eukaryotes, more commonly known as “protists.” The organisms on any one branch share a common ancestor. In this book, we adopt the convention that time flows from left to right, so this tree (and other trees in this book) lies on its side, with its root—the common ancestor—at the left.

yourBioPortal.com

GO TO Web Activity 1.2 • The Major Groups of Organisms

have yet to be discovered by humans (see Section 32.4 for a discussion of how we know this). When we encounter a new species, its placement on the tree of life immediately tells us a great deal about its biology. In addition, understanding relationships among species allows biologists to make predictions about species that have not yet been studied, based on our knowledge of those that have.

For example, until phylogenetic methods were developed, it took years of investigation to isolate and identify most newly encountered human pathogens, and even longer to discover how these pathogens moved into human populations. Today, pathogens that cause diseases such as the flu are identified quickly on the basis of their evolutionary relationships. Placement in an evolutionary tree also gives us clues about the disease's biology, possible effective treatments, and the origin of the pathogen (see Chapters 21 and 22).

1.2 RECAP

The first cellular life on Earth was prokaryotic and arose about 4 billion years ago. The complexity of the organisms that exist today is the result of several important evolutionary events, including the evolution of photosynthesis, eukaryotic cells, and multicellularity. The genetic relationships of all organisms can be shown as a branching tree of life.

- Discuss the evolutionary significance of photosynthesis. **See pp. 10–11**
- What do the domains of life represent? What are the major groups of eukaryotes? **See p. 12 and Figure 1.10**

In February of 1676, Robert Hooke received a letter from the physicist Sir Isaac Newton in which Newton famously re-

marked, "If I have seen a little further, it is by standing on the shoulders of giants." We all stand on the shoulders of giants, building on the research of earlier scientists. By the end of this course, you will know more about evolution than Darwin ever could have, and you will know infinitely more about cells than Schleiden and Schwann did. Let's look at the methods biologists use to expand our knowledge of life.

1.3 How Do Biologists Investigate Life?

Regardless of the many different tools and methods used in research, all scientific investigations are based on *observation* and *experimentation*. In both, scientists are guided by the *scientific method*, one of the most powerful tools of modern science.

Observation is an important skill

Biologists have always observed the world around them, but today our ability to observe is greatly enhanced by technologies such as electron microscopes, DNA chips, magnetic resonance imaging, and global positioning satellites. These technologies have improved our ability to observe at all levels, from the distribution of molecules in the body to the distribution of fish in the oceans. For example, not too long ago marine biologists were only able to observe the movement of fish in the ocean by putting physical tags on the fish, releasing them, and hoping that a fisherman would catch that fish and send back the tag—and even that would reveal only where the fish ended up. Today we can attach electronic recording devices to fish that continuously record not only where the fish is, but also how deep it swims and the temperature and salinity of the water around it (**Figure 1.11**). The tags download this information to a satellite, which relays it back to researchers. Suddenly we are acquiring a great deal of knowledge about the distribution of life in the oceans—information that is relevant to studies of climate change.

Technologies that enable us to *quantify* observations are very important in science. For example, for hundreds of years species were classified by generally qualitative descriptions of the physical differences between them. There was no way of objectively calculating evolutionary distances between organisms, and biologists had to depend on the fossil record for insight. Today our ability to rapidly analyze DNA sequences enables quantitative estimates of evolutionary distances, as described in Parts Five and Six of this book. The ability to gather quantitative observations adds greatly to the biologist's ability to make strong conclusions.

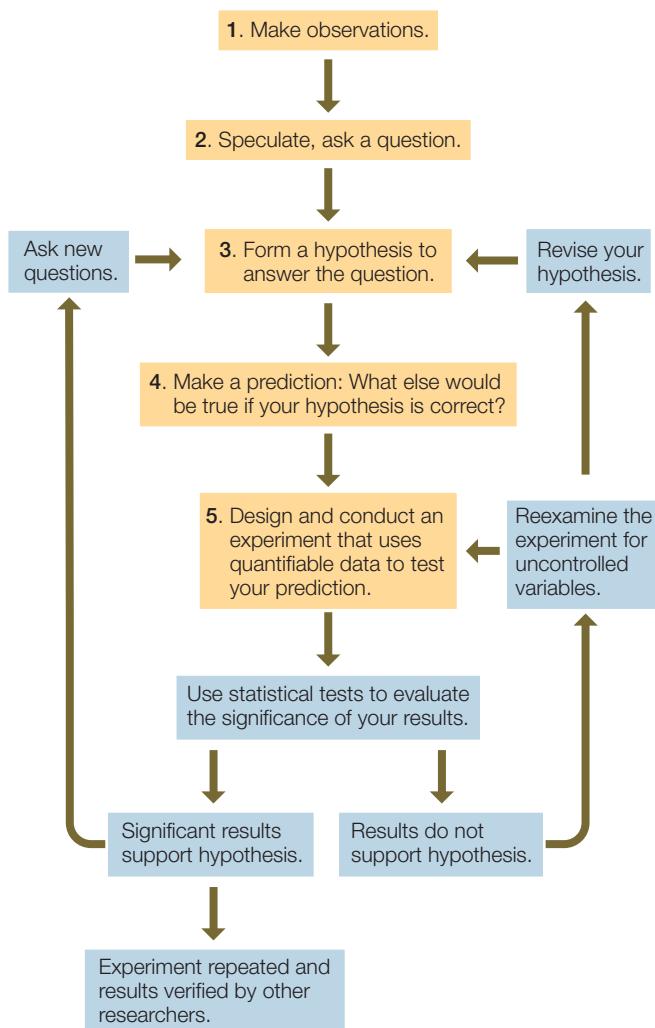


1.11 Tuna Tracking Marine biologist Barbara Block attaches computerized data recording tags (inset) to a live bluefin tuna before returning it to the ocean. Such tags make it possible to track an individual tuna wherever it travels in the world's oceans.

The scientific method combines observation and logic

Observations lead to questions, and scientists make additional observations and do experiments to answer those questions. The conceptual approach that underlies most modern scientific investigations is the **scientific method**. This powerful tool, also called the *hypothesis–prediction (H–P) method*, has five steps: (1) making *observations*; (2) asking *questions*; (3) forming *hypotheses*, or tentative answers to the questions; (4) making *predictions* based on the hypotheses; and (5) *testing* the predictions by making additional observations or conducting experiments (Figure 1.12).

After posing a question, a scientist uses *inductive logic* to propose a tentative answer. Inductive logic involves taking observations or facts and creating a new proposition that is compatible with those observations or facts. Such a tentative proposition is called a **hypothesis**. In formulating a hypothesis, scientists put together the facts they already know to formulate one or more possible answers to the question. For example, at the opening of



1.12 The Scientific Method The process of observation, speculation, hypothesis, prediction, and experimentation is the cornerstone of modern science. Answers gleaned through experimentation lead to new questions, more hypotheses, further experiments, and expanding knowledge.

this chapter you learned that scientists have observed the rapid decline of amphibian populations worldwide and are asking why. Some scientists have hypothesized that a fungal disease is a cause; other scientists have hypothesized that increased exposure to ultraviolet radiation is a cause. Tyrone Hayes hypothesized that exposure to agricultural chemicals could be a cause. He knew that the most widely used chemical herbicide is atrazine; that it is mostly applied in the spring, when amphibians are breeding; and that atrazine is a common contaminant in the waters in which amphibians live as they develop into adults.

The next step in the scientific method is to apply a different form of logic—*deductive logic*—to make predictions based on the hypothesis. Deductive logic starts with a statement believed to be true and then goes on to predict what facts would also have to be true to be compatible with that statement. Based on his hypothesis, Tyrone Hayes predicted that frog tadpoles exposed to atrazine would show adverse effects of the chemical once they reached adulthood.

Good experiments have the potential to falsify hypotheses

Once predictions are made from a hypothesis, experiments can be designed to test those predictions. The most informative experiments are those that have the ability to show that the prediction is wrong. If the prediction is wrong, the hypothesis must be questioned, modified, or rejected.

There are two general types of experiments, both of which compare data from different groups or samples. A *controlled experiment* manipulates one or more of the factors being tested; *comparative experiments* compare unmanipulated data gathered from different sources. As described at the opening of this chapter, Tyrone Hayes and his colleagues conducted both types of experiment to test the prediction that the herbicide atrazine, a contaminant in freshwater ponds and streams throughout the world, affects the development of frogs.

In a **controlled experiment**, we start with groups or samples that are as similar as possible. We predict on the basis of our hypothesis that some critical factor, or **variable**, has an effect on the phenomenon we are investigating. We devise some method to manipulate *only that variable* in an “experimental” group and compare the resulting data with data from an unmanipulated “control” group. If the predicted difference occurs, we then apply statistical tests to ascertain the probability that the manipulation created the difference (as opposed to the difference being the result of random chance). Figure 1.13 describes one of the many controlled experiments performed by the Hayes laboratory to quantify the effects of atrazine on male frogs.

The basis of controlled experiments is that one variable is manipulated while all others are held constant. The variable that is manipulated is called the *independent variable*, and the response that is measured is the *dependent variable*. A good controlled experiment is not easy to design because biological variables are so interrelated that it is difficult to alter just one.

A **comparative experiment** starts with the prediction that there will be a difference between samples or groups based on the hypothesis. In comparative experiments, however, we can-

INVESTIGATING LIFE

1.13 Controlled Experiments Manipulate a Variable

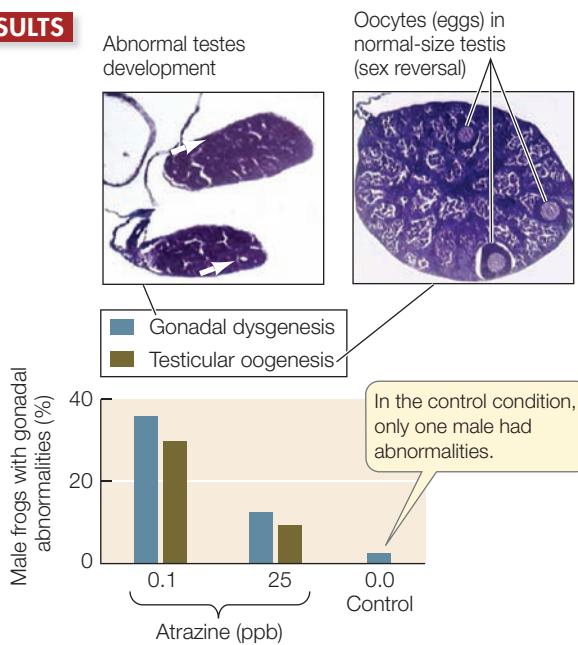
The Hayes laboratory created controlled environments that differed only in the concentrations of atrazine in the water. Eggs from leopard frogs (*Rana pipiens*) raised specifically for laboratory use were allowed to hatch and the tadpoles were separated into experimental tanks containing water with different concentrations of atrazine.

HYPOTHESIS Exposure to atrazine during larval development causes abnormalities in the reproductive system of male frogs.

METHOD

- Establish 9 tanks in which all attributes are held constant except the water's atrazine concentrations. Establish 3 atrazine conditions (3 replicate tanks per condition): 0 ppb (control condition), 0.1 ppb, and 25 ppb.
- Place *Rana pipiens* tadpoles from laboratory-reared eggs in the 9 tanks (30 tadpoles per replicate).
- When tadpoles have transitioned into adults, sacrifice the animals and evaluate their reproductive tissues.
- Test for correlation of degree of atrazine exposure with the presence of abnormalities in the reproductive systems of male frogs.

RESULTS



CONCLUSION Exposure to atrazine at concentrations as low as 0.1 ppb induces abnormalities in the male reproductive systems of frogs. The effect is not proportional to the level of exposure.

Go to yourBioPortal.com for original citations, discussions, and relevant links for all INVESTIGATING LIFE figures.

not control the variables; often we cannot even identify all the variables that are present. We are simply gathering and comparing data from different sample groups.

When his controlled experiments indicated that atrazine indeed affects reproductive development in frogs, Hayes and his colleagues performed a comparative experiment. They collected frogs and water samples from eight widely separated sites across the United States and compared the incidence of abnormal frogs from environments with very different levels of atrazine (Figure 1.14). Of course, the sample sites differed in many ways besides the level of atrazine present.

The results of experiments frequently reveal that the situation is more complex than the hypothesis anticipated, thus raising new questions. In the Hayes experiments, for example, there was no clear direct relationship between the *amount* of atrazine present and the percentage of abnormal frogs: there were fewer abnormal frogs at the highest concentrations of atrazine than at lower concentrations. There are no “final answers” in science. Investigations consistently reveal more complexity than we expect. The scientific method is a tool to identify, assess, and understand that complexity.

yourBioPortal.com

[GO TO](#) Animated Tutorial 1.1 • The Scientific Method

Statistical methods are essential scientific tools

Whether we do comparative or controlled experiments, at the end we have to decide whether there is a difference between the samples, individuals, groups, or populations in the study. How do we decide whether a measured difference is enough to support or falsify a hypothesis? In other words, how do we decide in an unbiased, objective way that the measured difference is significant?

Significance can be measured with statistical methods. Scientists use statistics because they recognize that variation is always present in any set of measurements. Statistical tests calculate the probability that the differences observed in an experiment could be due to random variation. The results of statistical tests are therefore probabilities. A statistical test starts with a **null hypothesis**—the premise that no difference exists. When quantified observations, or **data**, are collected, statistical methods are applied to those data to calculate the likelihood that the null hypothesis is correct.

More specifically, statistical methods tell us the probability of obtaining the same results by chance even if the null hypothesis were true. We need to eliminate, insofar as possible, the chance that any differences showing up in the data are merely the result of random variation in the samples tested. Scientists generally conclude that the differences they measure are significant if statistical tests show that the *probability of error* (that is, the probability that the same results can be obtained by mere chance) is 5 percent or lower.

Not all forms of inquiry are scientific

Science is a unique human endeavor that is bounded by certain standards of practice. Other areas of scholarship share with science the practice of making observations and asking ques-

INVESTIGATING LIFE

1.14 Comparative Experiments Look for Differences among Groups

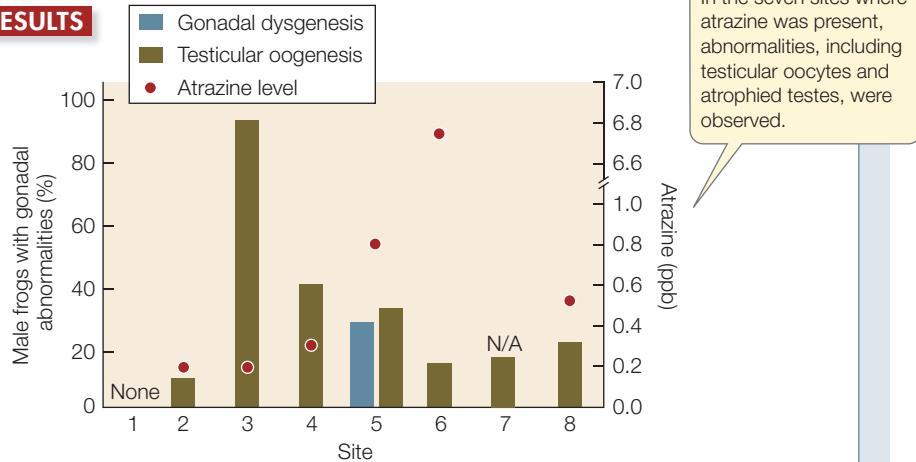
To see whether the presence of atrazine correlates with reproductive system abnormalities in male frogs, the Hayes lab collected frogs and water samples from different locations around the U.S. The analysis that followed was “blind,” meaning that the frogs and water samples were coded so that experimenters working with each specimen did not know which site the specimen came from.

HYPOTHESIS Presence of the herbicide atrazine in environmental water correlates with reproductive system abnormalities in frog populations.

METHOD

1. Based on commercial sales of atrazine, select 4 sites (sites 1–4) less likely and 4 sites (sites 5–8) more likely to be contaminated with atrazine.
2. Visit all sites in the spring (i.e., when frogs have transitioned from tadpoles into adults); collect frogs and water samples.
3. In the laboratory, sacrifice frogs and examine their reproductive tissues, documenting abnormalities.
4. Analyze the water samples for atrazine concentration (the sample for site 7 was not tested).
5. Quantify and correlate the incidence of reproductive abnormalities with environmental atrazine concentrations.

RESULTS



CONCLUSION Reproductive abnormalities exist in frogs from environments in which aqueous atrazine concentration is 0.2 ppb or above. The incidence of abnormalities does not appear to be proportional to atrazine concentration at the time of transition to adulthood.

FURTHER INVESTIGATION: The highest proportion of abnormal frogs was found at site 3, located on a wildlife reserve in Wyoming. What kind of data and observations would you need to suggest possible explanations for this extremely high incidence?

Go to yourBioPortal.com for original citations, discussions, and relevant links for all INVESTIGATING LIFE figures.

tions, but scientists are distinguished by what they do with their observations and how they answer their questions. Data, subjected to appropriate statistical analysis, are critical in the testing of hypotheses. The scientific method is the most powerful way humans have devised for learning about the world and how it works.

Scientific explanations for natural processes are objective and reliable because the hypotheses proposed *must be testable* and *must have the potential of being rejected* by direct observations and experiments. Scientists must clearly describe the methods they use to test hypotheses so that other scientists can repeat their results. Not all experiments are repeated, but surprising or controversial results are always subjected to independent verification. Scientists worldwide share this process of testing and rejecting hypotheses, contributing to a common body of scientific knowledge.

If you understand the methods of science, you can distinguish science from non-science. Art, music, and literature all contribute to the quality of human life, but they are not science. They do not use the scientific method to establish what is fact. Religion is not science, although religions have historically purported to explain natural events ranging from unusual weather patterns to crop failures to human diseases. Most such phenomena that at one time were mysterious can now be explained in terms of scientific principles.

The power of science derives from the uncompromising objectivity and absolute dependence on evidence that comes from *reproducible and quantifiable observations*. A religious or spiritual explanation of a natural phenomenon may be coherent and satisfying for the person holding that view, but it is not testable, and therefore it is not science. To invoke a supernatural explanation (such as a “creator” or “intelligent designer” with no known bounds) is to depart from the world of science.

Science describes the facts about how the world works, not how it “ought to be.” Many scientific advances that have contributed to human welfare have also raised major ethical issues. Recent developments in genetics and developmental biology, for example, enable us to select the sex of our children, to use stem cells to repair our bodies, and to modify the

human genome. Although scientific knowledge allows us to do these things, science cannot tell us whether or not we should do them, or, if we choose to do so, how we should regulate them.

To make wise decisions about public policy, we need to employ the best possible ethical reasoning in deciding which outcomes we should strive for.

1.3 RECAP

The scientific method of inquiry starts with the formulation of hypotheses based on observations and data. Comparative and controlled experiments are carried out to test hypotheses.

- Explain the relationship between a hypothesis and an experiment. **See p. 14 and Figure 1.12**
- What is controlled in a controlled experiment? **See p. 14 and Figure 1.13**
- What features characterize questions that can be answered only by using a comparative approach? **See pp. 14–15 and Figure 1.14**
- Do you understand why arguments must be supported by quantifiable and reproducible data in order to be considered scientific? **See pp. 15–16**

The vast scientific knowledge accumulated over centuries of human civilization allows us to understand and manipulate aspects of the natural world in ways that no other species can. These abilities present us with challenges, opportunities, and above all, responsibilities.

1.4 How Does Biology Influence Public Policy?

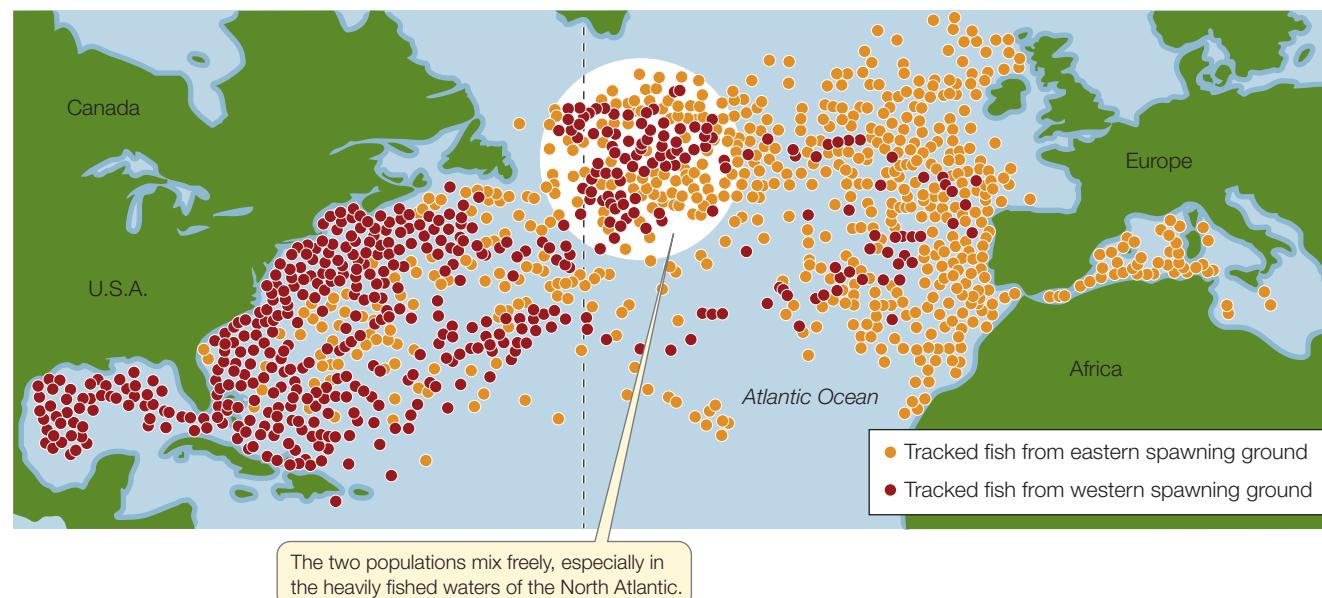
Agriculture and medicine are two important human activities that depend on biological knowledge. Our ancestors unknowingly applied the principles of evolutionary biology when they domesticated plants and animals, and people have speculated about the causes of diseases and searched for methods to combat them since ancient times. Long before the microbial causes of diseases were known, people recognized that infections could

be passed from one person to another, and the isolation of infected persons has been practiced as long as written records have been available.

Today, thanks to the deciphering of genomes and our newfound ability to manipulate them, vast new possibilities exist for controlling human diseases and increasing agricultural productivity, but these capabilities raise ethical and policy issues. How much and in what ways should we tinker with the genes of humans and other species? Does it matter whether the genomes of our crop plants and domesticated animals are changed by traditional methods of controlled breeding and crossbreeding or by the biotechnology of gene transfer? What rules should govern the release of genetically modified organisms into the environment? Science alone cannot provide all the answers, but wise policy decisions must be based on accurate scientific information.

Biologists are increasingly called on to advise government agencies concerning the laws, rules, and regulations by which society deals with the increasing number of challenges that have at least a partial biological basis. As an example of the value of scientific knowledge for the assessment and formulation of public policy, let's return to the tracking study of bluefin tuna introduced in Section 1.3. Prior to this study, both scientists and fishermen knew that bluefins had a western breeding ground in the Gulf of Mexico and an eastern breeding ground in the Mediterranean Sea (**Figure 1.15**). Overfishing had led to declining numbers of fish in the western-breeding populations, to the point of these populations being endangered.

1.15 Bluefin Tuna Do Not Recognize Boundaries It was assumed that tuna from western-breeding populations and those from eastern-breeding populations also fed on their respective sides of the Atlantic, so separate fishing quotas were established on either side of 45° W longitude (dashed line) to allow the endangered western population to recover. However, tracking data shows that the two populations *do not* remain separate after spawning, so in fact the established policy does not protect the western population.



Initially it was assumed by scientists, fishermen, and policy makers alike that the eastern and western populations had geographically separate feeding grounds as well as separate breeding grounds. Acting on this assumption, an international commission drew a line down the middle of the Atlantic Ocean and established strict fishing quotas on the western side of the line, with the intent of allowing the western population to recover. New tracking data, however, revealed that in fact the eastern and western bluefin populations mix freely on their feeding grounds across the entire North Atlantic—a swath of ocean that includes the most heavily fished waters in the world. Tuna caught on the eastern side of the line could just as likely be from the western breeding population as the eastern; thus the established policy was not achieving its intended goal.

Policy makers take more things into consideration than scientific knowledge and recommendations. For example, studies on the effects of atrazine on amphibians have led one U.S. group, the Natural Resources Defense Council, to take legal action to have atrazine banned on the basis of the Endangered Species Act. The U.S. Environmental Protection Agency, however, must also consider the potential loss to agriculture that such a ban would create and has continued to approve atrazine's use as long as environmental levels do not exceed 30 to 40 ppb—which is 300 to 400 times the levels shown to induce abnormalities in the Hayes studies. Scientific conclusions do not always prevail in the political world.

Another reason for studying biology is to understand the effects of the vastly increased human population on its environment. Our use of natural resources is putting stress on the ability of Earth's ecosystems to continue to produce the goods and services on which our society depends. Human activities are changing global climates, causing the extinctions of a large number of species like the amphibians featured in this chapter, and spreading new diseases while facilitating the resurgence of old ones. The rapid spread of flu viruses has been facilitated by modern modes of transportation, and the recent resurgence of tuberculosis is the result of the evolution of bacteria that are resistant to antibiotics. Biological knowledge is vital for determining the causes of these changes and for devising wise policies to deal with them.

Beyond issues of policy and pragmatism lies the human “need to know.” Human beings are fascinated by the richness and diversity of life, and most people want to know more about organisms and how they interact. Human curiosity might even be seen as an adaptive trait—it is possible that such a trait could have been selected for if individuals who were motivated to learn about their surroundings were likely to have survived and reproduced better, on average, than their less curious relatives. Far from ending the process, new discoveries and greater knowledge typically engender questions no one thought to ask before. There are vast numbers of questions for which we do not yet have answers, and the most important motivator of most scientists is curiosity.

CHAPTER SUMMARY

1.1 What Is Biology?

- **Biology** is the scientific study of living organisms, including their characteristics, functions, and interactions. Cells are the basic structural and physiological units of life. The **cell theory** states that all life consists of cells and that all cells come from preexisting cells.
- All living organisms are related to one another through descent with modification. **Evolution** by **natural selection** is responsible for the diversity of **adaptations** found in living organisms.
- The instructions for a cell are contained in its **genome**, which consists of **DNA** molecules made up of sequences of **nucleotides**. Specific segments of DNA called **genes** contain the information the cell uses to make **proteins**. **Review Figure 1.4**
- Living organisms regulate their internal environment. They also interact with other organisms of the same and different species. Biologists study life at all these levels of organization. **Review Figure 1.6**, **WEB ACTIVITY 1.1**
- Biological knowledge obtained from a **model system** may be generalized to other species.

1.2 How Is All Life on Earth Related?

- Biologists use fossils, anatomical similarities and differences, and molecular comparisons of genomes to reconstruct the history of life. **Review Figure 1.8**

- Life first arose by chemical evolution. Cells arose early in the evolution of life.
- **Photosynthesis** was an important evolutionary step because it changed Earth's atmosphere and provided a means of capturing energy from sunlight.
- The earliest organisms were **prokaryotes**. Organisms called **eukaryotes**, with more complex cells, arose later. Eukaryotic cells have discrete intracellular compartments, called **organelles**, including a nucleus that contains the cell's genetic material.
- The genetic relationships of **species** can be represented as an evolutionary tree. Species are grouped into three **domains**: **Archaea**, **Bacteria**, and **Eukarya**. Archaea and Bacteria are domains of unicellular prokaryotes. Eukarya contains diverse groups of protists (most but not all of which are unicellular) and the multicellular plants, fungi, and animals. **Review Figure 1.10**, **WEB ACTIVITY 1.2**

1.3 How Do Biologists Investigate Life?

- The **scientific method** used in most biological investigations involves five steps: making observations, asking questions, forming hypotheses, making predictions, and testing those predictions. **Review Figure 1.12**
- **Hypotheses** are tentative answers to questions. Predictions made on the basis of a hypothesis are tested with additional

observations and two kinds of **experiments: comparative** and **controlled experiments**. Review Figures 1.13 and 1.14, ANIMATED TUTORIAL 1.1

- Statistical methods are applied to **data** to establish whether or not the differences observed are significant or whether they could be the result of chance. These methods start with the **null hypothesis** that there are no differences.

- Science can tell us how the world works, but it cannot tell us what we should or should not do.

1.4 How Does Biology Influence Public Policy?

- Biologists are often called on to advise government agencies on the solution of important problems that have a biological component.

FOR DISCUSSION

1. Even if we knew the sequences of all of the genes of a single-celled organism and could cause those genes to be expressed in a test tube, it would still be incredibly difficult to create a functioning organism. Why do you think this is so? In light of this fact, what do you think of the statement that the genome contains all of the information for a species?
2. Why is it so important in science that we design and perform tests capable of falsifying a hypothesis?
3. What features characterize questions that can be answered only by using a comparative approach?
4. Cite an example of how you apply aspects of the scientific method to solve problems in your daily life.

ADDITIONAL INVESTIGATION

1. The abnormalities of frogs in Tyrone Hayes's studies were associated with the presence of a herbicide in the environment. That herbicide did not kill the frogs, but it feminized the males. How would you investigate whether this effect could lead to decreased reproductive capacity for the frog populations in nature?
2. Just as all cells come from preexisting cells, all mitochondria—the cell organelles that convert energy in food to a form of energy that can do biological work—come from preexisting mitochondria. Cells do not synthesize mitochondria from the genetic information in their nuclei. What investigations would you carry out to understand the nature of mitochondria?

WORKING WITH DATA (GO TO [yourBioPortal.com](#))

Feminization of Frogs Analogous to the experiment shown in Figure 1.13, this exercise asks you to graph data about the size of the laryngeal (throat) muscles required to produce male mating calls in the frog *Xenopus laevis*. After plotting data from

frogs exposed to different levels of the herbicide atrazine during their development, you will formulate conclusions about the effects of the herbicide on this physical attribute and speculate about what these effects might mean.

2

Small Molecules and the Chemistry of Life

A hairy story

"**Y**ou are what you eat—and that is recorded in your hair." Two scientists at the University of Utah are responsible for adding the last phrase to this famous saying about body chemistry. Ecologist Jim Ehleringer and chemist Thure Cerling showed that the composition of human hair reflects the region where a person lives.

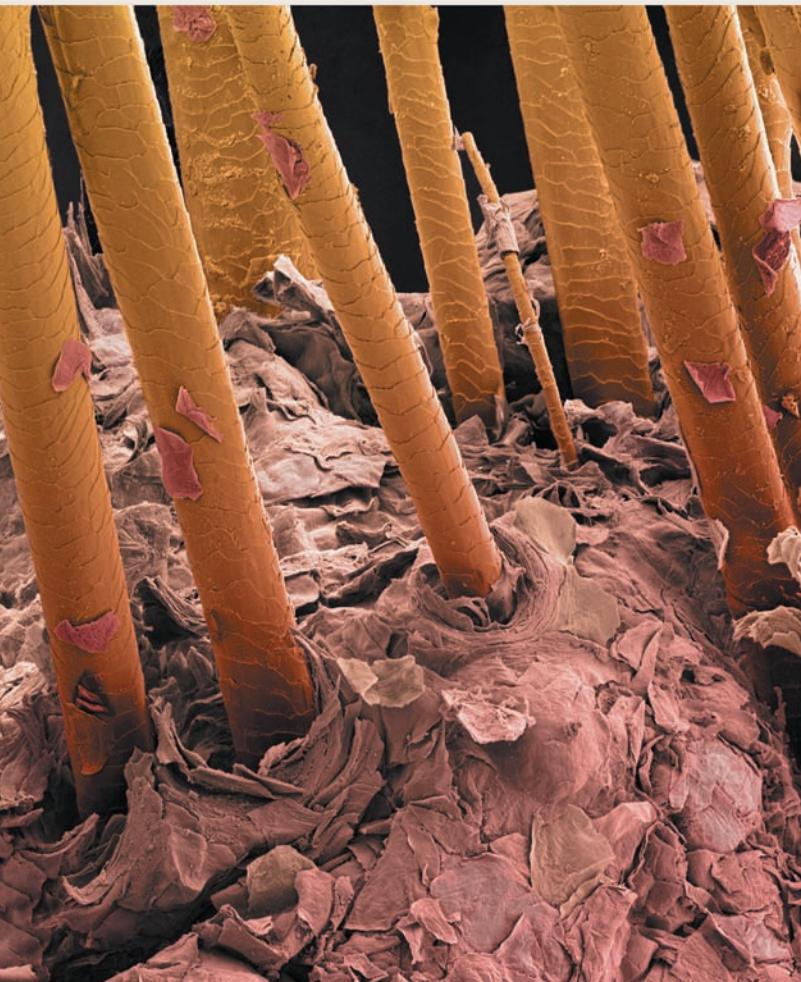
As we pointed out in Chapter 1, living things are made up of the same kinds of atoms that make up the inanimate universe. Two of those atoms are hydrogen (H) and

oxygen (O), which combine to form water (H_2O). Both atoms have naturally occurring variants called *isotopes*, which have the same chemical properties but different weights because their nuclei have different numbers of particles called *neutrons*.

When water evaporates from the ocean, it forms clouds that move inland and release rain. Water made up of the heavier H and O isotopes is heavier and tends to fall more readily than water containing the lighter isotopes. Warm rains tend to be heavier than cooler precipitation. People living on the coast or in regions where there are frequent warm rains consume heavier water and foods made from water than people living in cooler, inland areas (assuming, of course, that their beverages and produce come from the same area they live in). And, since you are what you eat, the heavy H and O atoms become part of their bodies.

Our hair contains abundant H and O atoms, many obtained from local water. Ehleringer and Cerling wondered whether the ratios of heavy-to-light H and O in hair reflected the ratio of heavy-to-light H_2O in the local water. To address this question, Ehleringer's wife and Cerling's children and their friends went on a hair-collecting trip across the United States, collecting hair trimmings from barbershop floors while at the same time filling test tubes with local water. Back at the lab, scientists tested the samples and found that the ratios of heavy to light isotopes in the hair did indeed reflect these same ratios in the local water.

While this information is intrinsically fascinating, it is also potentially useful. For example, police could use hair analysis to evaluate a suspect's alibi: "You say you've been in Montana for the past month? Your hair sample indicates that you were in a warm coastal area." Such conflicting evidence could form the basis of further investigation.



Hair Tells a Tale The ratio in hair protein of the heavy isotope ^{18}O to its lighter counterpart ^{16}O reflects the ratios in local water.



Free Samples Need hair samples for a research project? Try the local barber shop.

Or anthropologists might analyze hair samples from graves to work out migration patterns of human groups.

The understanding that life is based on chemistry and obeys universal laws of chemistry and physics is relatively new in human history. Until the nineteenth century, a “vital force” (from the Latin *vitalis*, “of life”) was presumed to be responsible for life. This vital force was seen as distinct from the mechanistic forces governing physics and chemistry. Many people still assume that a vital force exists, but the physical–chemical view of life has led to great advances in biological science and is the cornerstone of modern medicine and agriculture.

IN THIS CHAPTER we will introduce the constituents of matter: atoms, their variety, their properties, and their capacity to combine with other atoms. We will consider how matter changes, including changes in state (solid to liquid to gas), and changes caused by chemical reactions. We will examine the structure and properties of water and its relationship to chemical acids and bases.

CHAPTER OUTLINE

- 2.1 How Does Atomic Structure Explain the Properties of Matter?**
- 2.2 How Do Atoms Bond to Form Molecules?**
- 2.3 How Do Atoms Change Partners in Chemical Reactions?**
- 2.4 What Makes Water So Important for Life?**

2.1 How Does Atomic Structure Explain the Properties of Matter?

All matter is composed of **atoms**. Atoms are tiny—more than a trillion (10^{12}) of them could fit on top of the period at the end of this sentence. Each atom consists of a dense, positively charged **nucleus**, around which one or more negatively charged **electrons** move (Figure 2.1). The nucleus contains one or more positively charged **protons** and may contain one or more **neutrons** with no electrical charge. Atoms and their component particles have volume and mass, which are characteristics of all matter. *Mass* is a measure of the quantity of matter present; the greater the mass, the greater the quantity of matter.

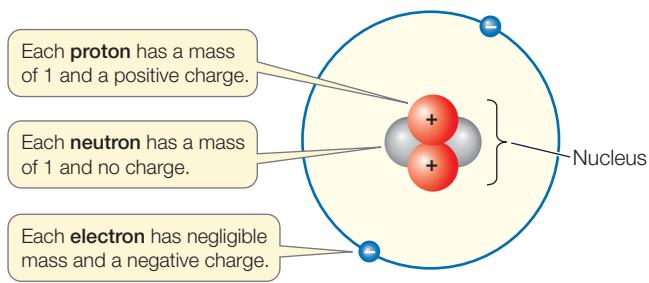
The mass of a proton serves as a standard unit of measure called the *dalton* (named after the English chemist John Dalton) or **atomic mass unit (amu)**. A single proton or neutron has a mass of about 1 dalton (Da), which is 1.7×10^{-24} grams ($0.00000000000000000000000000000017$ g). That’s tiny, but an electron is even tinier at 9×10^{-28} g (0.0005 Da). Because the mass of an electron is negligible compared with the mass of a proton or a neutron, the contribution of electrons to the mass of an atom can usually be ignored when measurements and calculations are made. It is electrons, however, that determine how atoms will combine with other atoms to form stable associations.

Each proton has a positive electric charge, defined as +1 unit of charge. An electron has a negative charge equal and opposite to that of a proton (-1). The neutron, as its name suggests, is electrically neutral, so its charge is 0. Charges that are different (+/-) attract each other, whereas charges that are alike (+/+, -/-) repel each other. Atoms are electrically neutral because the number of electrons in an atom equals the number of protons.

An element consists of only one kind of atom

An **element** is a pure substance that contains only one kind of atom. The element hydrogen consists only of hydrogen atoms; the element iron consists only of iron atoms. The atoms of each element have certain characteristics or properties that distinguish them from the atoms of other elements. These properties include their mass and how they interact and associate with other atoms.

The more than 100 elements found in the universe are arranged in the *periodic table* (Figure 2.2). Each element has its own one- or two-letter chemical symbol. For example, H stands for hydrogen, C for carbon, and O for oxygen. Some symbols come from other languages: Fe (from the Latin, *ferrum*) stands for iron, Na (Latin, *natrium*) for sodium, and W (German, *wolfram*) for tungsten.



2.1 The Helium Atom This representation of a helium atom is called a Bohr model. It exaggerates the space occupied by the nucleus. In reality, although the nucleus accounts for virtually all of the atomic mass, it occupies only about 1/10,000 of the atom's volume. The Bohr model is also inaccurate in that it represents the electron as a discrete particle in a defined orbit around the nucleus.

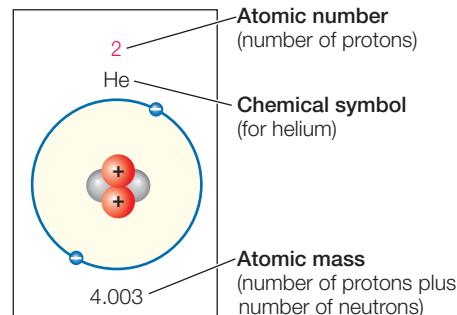
The elements of the periodic table are not found in equal amounts. Stars have abundant amounts of hydrogen and helium. Earth's crust, and the surfaces of the neighboring planets, are almost half oxygen, 28 percent silicon, 8 percent aluminum, and between 2 and 5 percent each of sodium, magnesium, potassium, calcium, and iron. They contain much smaller amounts of the other elements.

About 98 percent of the mass of every living organism (bacterium, turnip, or human) is composed of just six elements: carbon,

bon, hydrogen, nitrogen, oxygen, phosphorus, and sulfur. The chemistry of these six elements will be our primary concern in this chapter, but other elements found in living organisms are important as well. Sodium and potassium, for example, are essential for nerve function; calcium can act as a biological signal; iodine is a component of a vital hormone; and magnesium is bound to chlorophyll in plants. The physical and chemical (reactive) properties of atoms depend on the number of subatomic particles they contain.

Each element has a different number of protons

An element differs from other elements by the number of protons in the nucleus of each of its atoms; the number of protons is designated the **atomic number**. This atomic number is unique



1 H 1.0079	elements according to their physical and chemical properties. Elements 1–92 occur in nature; elements with atomic numbers above 92 were created in the laboratory.												The six elements highlighted in yellow make up 98% of the mass of most living organisms.							
3 Li 6.941	4 Be 9.012	Elements in the same vertical columns have similar properties because they have the same number of electrons in their outermost shell.						Elements highlighted in orange are present in small amounts in many organisms.						5 B 10.81	6 C 12.011	7 N 14.007	8 O 15.999	9 F 18.998	10 Ne 20.179	
11 Na 22.990	12 Mg 24.305														13 Al 26.982	14 Si 28.086	15 P 30.974	16 S 32.06	17 Cl 35.453	18 Ar 39.948
19 K 39.098	20 Ca 40.08	21 Sc 44.956	22 Ti 47.88	23 V 50.942	24 Cr 51.996	25 Mn 54.938	26 Fe 55.847	27 Co 58.933	28 Ni 58.69	29 Cu 63.546	30 Zn 65.38	31 Ga 69.72	32 Ge 72.59	33 As 74.922	34 Se 78.96	35 Br 79.909	36 Kr 83.80			
37 Rb 85.4778	38 Sr 87.62	39 Y 88.906	40 Zr 91.22	41 Nb 92.906	42 Mo 95.94	43 Tc (99)	44 Ru 101.07	45 Rh 102.906	46 Pd 106.4	47 Ag 107.870	48 Cd 112.41	49 In 114.82	50 Sn 118.69	51 Sb 121.75	52 Te 127.60	53 I 126.904	54 Xe 131.30			
55 Cs 132.905	56 Ba 137.34	71 Lu 174.97	72 Hf 178.49	73 Ta 180.948	74 W 183.85	75 Re 186.207	76 Os 190.2	77 Ir 192.2	78 Pt 195.08	79 Au 196.967	80 Hg 200.59	81 Tl 204.37	82 Pb 207.19	83 Bi 208.980	84 Po (209)	85 At (210)	86 Rn (222)			
87 Fr (223)	88 Ra (226)	103 Lr (260)	104 Rf (261)	105 Db (262)	106 Sg (266)	107 Bh (264)	108 Hs (269)	109 Mt (268)	110 (269)	111 (272)	112 (277)	113 (285)	114 (289)	115 (289)	116 (289)	117 (293)	118 (293)			
Masses in parentheses indicate unstable elements that decay rapidly to form other elements.															Elements without a chemical symbol are as yet unnamed.					
Lanthanide series		57 La 138.906	58 Ce 140.12	59 Pr 140.9077	60 Nd 144.24	61 Pm (145)	62 Sm 150.36	63 Eu 151.96	64 Gd 157.25	65 Tb 158.924	66 Dy 162.50	67 Ho 164.930	68 Er 167.26	69 Tm 168.934	70 Yb 173.04					
Actinide series		89 Ac 227.028	90 Th 232.038	91 Pa 231.0359	92 U 238.02	93 Np 237.0482	94 Pu (244)	95 Am (243)	96 Cm (247)	97 Bk (247)	98 Cf (251)	99 Es (252)	100 Fm (257)	101 Md (258)	102 No (259)					

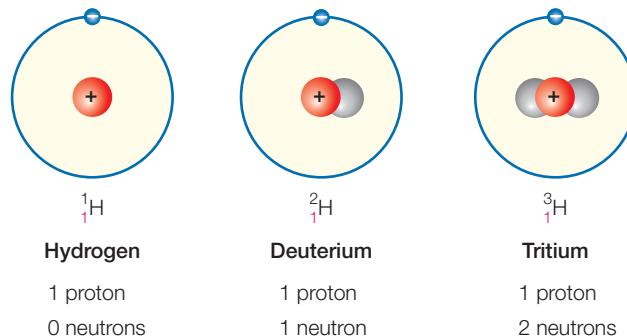
to each element and does not change. The atomic number of helium is 2, and an atom of helium always has two protons; the atomic number of oxygen is 8, and an atom of oxygen always has eight protons.

Along with a definitive number of protons, every element except hydrogen has one or more neutrons in its nucleus. The **mass number** of an atom is the total number of protons and neutrons in its nucleus. The nucleus of a carbon atom contains six protons and six neutrons, and has a mass number of 12. Oxygen has eight protons and eight neutrons, and has a mass number of 16. The mass number is essentially the mass of the atom in daltons (see below).

By convention, we often print the symbol for an element with the atomic number at the lower left and the mass number at the upper left, both immediately preceding the symbol. Thus hydrogen, carbon, and oxygen can be written as ^1_1H , $^{12}_6\text{C}$, and $^{16}_8\text{O}$, respectively.

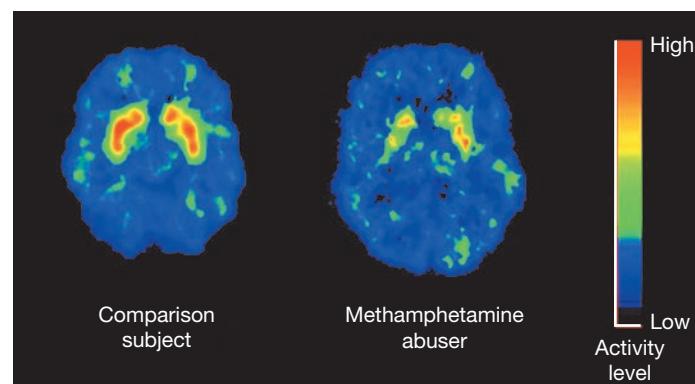
The number of neutrons differs among isotopes

In some elements, the number of neutrons in the atomic nucleus is not always the same. Different **isotopes** of the same element have the same number of protons, but different numbers of neutrons. Many elements have several isotopes. The isotopes of hydrogen shown below have special names, but the isotopes of most elements do not have distinct names.



The natural isotopes of carbon, for example, are ^{12}C (six neutrons in the nucleus), ^{13}C (seven neutrons), and ^{14}C (eight neutrons). Note that all three (called “carbon-12,” “carbon-13,” and “carbon-14”) have six protons, so they are all carbon. Most carbon atoms are ^{12}C , about 1.1 percent are ^{13}C , and a tiny fraction are ^{14}C . But all have virtually the same chemical reactivity, which is an important property for their use in experimental biology and medicine. An element’s **atomic weight** (or atomic mass) is the average of the mass numbers of a representative sample of atoms of that element, with all the isotopes in their normally occurring proportions. The atomic weight of carbon, taking into account all of its isotopes and their abundances, is thus 12.011. The fractional atomic weight results from averaging the contributing weights of all of the isotopes.

Most isotopes are stable. But some, called **radioisotopes**, are unstable and spontaneously give off energy in the form of α (alpha), β (beta), or γ (gamma) radiation from the atomic nucleus. Known as *radioactive decay*, this release of energy transforms the original atom. The type of transformation varies depending on



2.3 Tagging the Brain In these images from live people, a radioactive-labeled sugar is used to detect differences between the brain activity of a healthy person and that of a person who abuses methamphetamines. The more active a brain region is, the more sugar it takes up. The healthy brain (left) shows more activity in the region involved in memory (the red area) than the drug abuser’s brain does.

the radioisotope, but some can change the number of protons, so that the original atom becomes a different element.

With sensitive instruments, scientists can use the released radiation to detect the presence of radioisotopes. For instance, if an earthworm is given food containing a radioisotope, its path through the soil can be followed using a simple detector called a Geiger counter. Most atoms in living organisms are organized into stable associations called **molecules**. If a radioisotope is incorporated into a molecule, it acts as a tag or label, allowing researchers or physicians to trace the molecule in an experiment or in the body (Figure 2.3). Radioisotopes are also used to date fossils, an application described in Section 25.1.

Although radioisotopes are useful in research and in medicine, even a low dose of the radiation they emit has the potential to damage molecules and cells. However, these damaging effects are sometimes used to our advantage; for example, the radiation from ^{60}Co (cobalt-60) is used in medicine to kill cancer cells.

The behavior of electrons determines chemical bonding and geometry

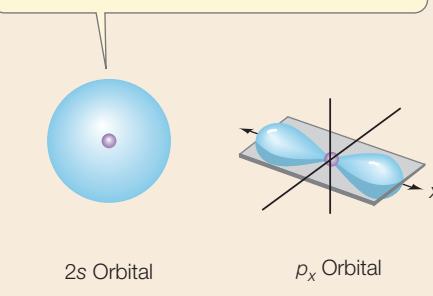
The characteristic number of electrons in an atom determines how it will combine with other atoms. Biologists are interested in how chemical changes take place in living cells. When considering atoms, they are concerned primarily with electrons because the behavior of electrons explains how chemical *reactions* occur. Chemical reactions alter the atomic compositions of substances and thus alter their properties. Reactions usually involve changes in the distribution of electrons between atoms.

The location of a given electron in an atom at any given time is impossible to determine. We can only describe a volume of space within the atom where the electron is likely to be. The region of space where the electron is found at least 90 percent of the time is the electron’s **orbital**. Orbitals have characteristic shapes and orientations, and a given orbital can be occupied by

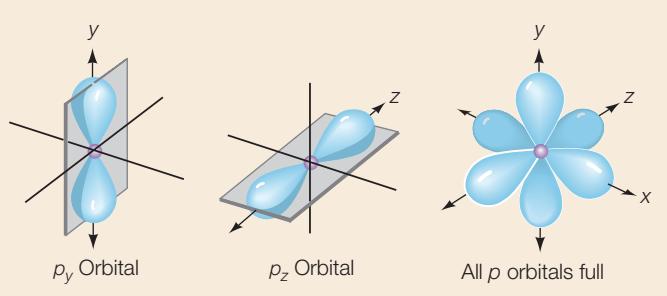
First shell:
The two electrons closest to the nucleus move in a spherical s orbital.



Second shell:
Two electrons occupy the 2s orbital, one of four orbitals in the second shell of electrons. The second shell can hold a total of eight electrons.



2.4 Electron Shells and Orbitals Each orbital holds a maximum of two electrons. The s orbitals have a lower energy level and fill with electrons before the p orbitals do.



yourBioPortal.com

GO TO Web Activity 2.1 • Electron Orbitals

Two electrons form a dumbbell-shaped x axis (p_x) orbital...

...two more fill the p_y orbital...

...and two fill the p_z orbital.

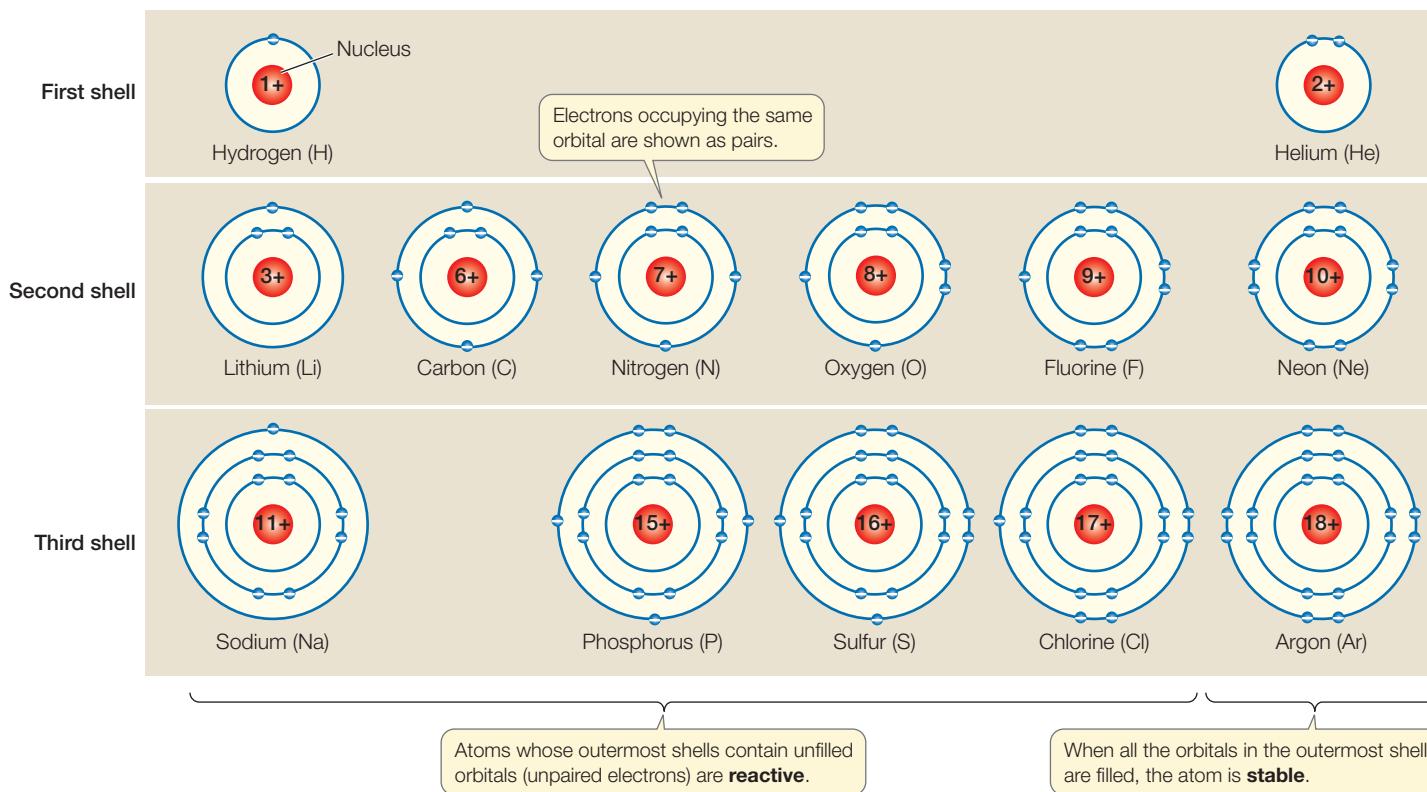
Six electrons fill all three p orbitals.

a maximum of two electrons (Figure 2.4). Thus any atom larger than helium (atomic number 2) must have electrons in two or more orbitals. As we move from lighter to heavier atoms in the periodic chart, the orbitals are filled in a specific sequence, in a series of what are known as **electron shells**, or *energy levels*, around the nucleus.

- **First shell:** The innermost electron shell consists of just one orbital, called an s orbital. A hydrogen atom (${}_1\text{H}$) has one electron in its first shell; helium (${}_2\text{He}$) has two. Atoms of all other elements have two or more shells to accommodate orbitals for additional electrons.

- **Second shell:** The second shell contains four orbitals (an s orbital and three p orbitals), and hence holds up to eight electrons. As depicted in Figure 2.4, the s orbitals have the shape of a sphere, while the p orbitals are directed at right angles to one another. The orientations of these orbitals in space contribute to the three-dimensional shapes of molecules when atoms link to other atoms.

2.5 Electron Shells Determine the Reactivity of Atoms Each shell can hold a specific maximum number of electrons. Each shell must be filled before electrons can occupy the next shell. The energy level of an electron is higher in a shell farther from the nucleus. An atom with unpaired electrons in its outermost shell can react (bond) with other atoms.



- **Additional shells:** Elements with more than ten electrons have three or more electron shells. The farther a shell is from the nucleus, the higher the energy level is for an electron occupying that shell.

The *s* orbitals fill with electrons first, and their electrons have the lowest energy level. Subsequent shells have different numbers of orbitals, but the outermost shells usually hold only eight electrons. In any atom, the outermost electron shell (the *valence shell*) determines how the atom combines with other atoms—that is, how the atom behaves chemically. When a valence shell with four orbitals contains eight electrons, there are no unpaired electrons, and the atom is *stable*—it will not react with other atoms (Figure 2.5). Examples of chemically stable elements are helium, neon, and argon. On the other hand, atoms that have one or more unpaired electrons in their outer shells are capable of reacting with other atoms.

Atoms with unpaired electrons (i.e., partially filled orbitals) in their outermost electron shells are unstable, and will undergo reactions in order to fill their outermost shells. Reactive atoms can attain stability either by sharing electrons with other atoms or by losing or gaining one or more electrons. In either case, the atoms involved are *bonded* together into stable associations called molecules. The tendency of atoms to form stable molecules so that they have eight electrons in their outermost shells is known as the *octet rule*. Many atoms in biologically important molecules—for example, carbon (C) and nitrogen (N)—follow this rule. An important exception is hydrogen (H), which attains stability when two electrons occupy its single shell (consisting of just one *s* orbital).

2.1 RECAP

The living world is composed of the same set of chemical elements as the rest of the universe. An atom consists of a nucleus of protons and neutrons, and a characteristic configuration of electrons in orbitals around the nucleus. This structure determines the atom's chemical properties.

- Describe the arrangement of protons, neutrons, and electrons in an atom. **See Figure 2.1**
- Use the periodic table to identify some of the similarities and differences in atomic structure among different elements (for example, oxygen, carbon, and helium). How does the configuration of the valence shell influence the placement of an element in the periodic table? **See p. 25 and Figures 2.2 and 2.5**
- How does bonding help a reactive atom achieve stability? **See p. 25 and Figure 2.5**

We have introduced the individual players on the biochemical stage—the atoms. We have shown how the energy levels of electrons drive an atomic “quest for stability.” Next we will describe the different types of chemical bonds that can lead to stability, joining atoms together into molecular structures with hosts of different properties.

2.2 How Do Atoms Bond to Form Molecules?

A **chemical bond** is an attractive force that links two atoms together in a molecule. There are several kinds of chemical bonds (Table 2.1). In this section we will begin with *covalent bonds*, the strong bonds that result from the sharing of electrons. Next we will examine *ionic bonds*, which form when an atom gains or loses one or more electrons to achieve stability. We will then consider other, weaker, kinds of interactions, including hydrogen bonds, which are enormously important to biology.

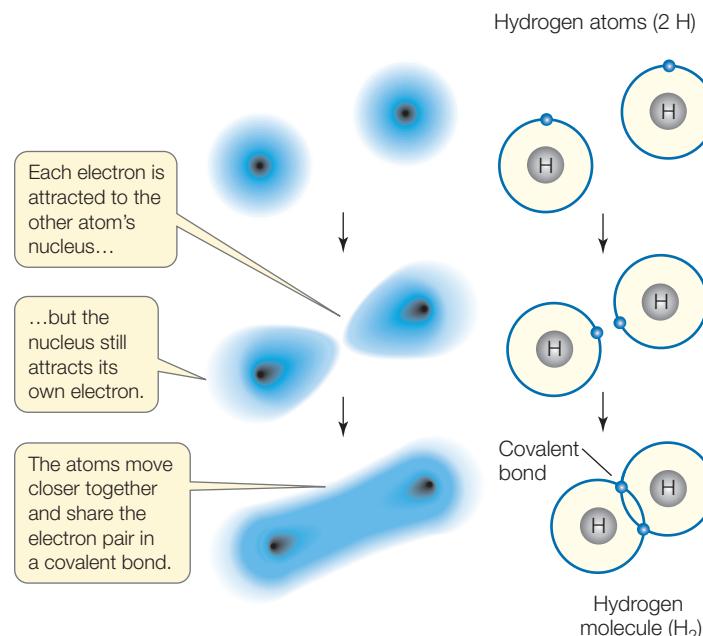
yourBioPortal.com

GO TO Animated Tutorial 2.1 • Chemical Bond Formation

Covalent bonds consist of shared pairs of electrons

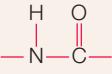
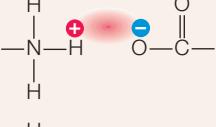
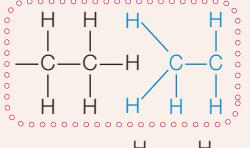
A **covalent bond** forms when two atoms attain stable electron numbers in their outermost shells by *sharing* one or more pairs of electrons. Consider two hydrogen atoms coming into close proximity, each with an unpaired electron in its single shell (Figure 2.6). When the electrons pair up, a stable association is formed, and this links the two hydrogen atoms in a covalent bond, resulting in H₂.

A **compound** is a substance made up of molecules with two or more elements bonded together in a fixed ratio. Methane gas (CH₄), water (H₂O), and table sugar (sucrose, C₁₂H₂₂O₁₁) are examples of compounds. The chemical symbols identify the different elements in a compound, and the subscript numbers indicate how many atoms of each element are present. Every compound has a **molecular weight** (molecular mass) that is the



2.6 Electrons Are Shared in Covalent Bonds Two hydrogen atoms can combine to form a hydrogen molecule. A covalent bond forms when the electron orbitals of the two atoms overlap in an energetically stable manner.

TABLE 2.1
Chemical Bonds and Interactions

NAME	BASIS OF INTERACTION	STRUCTURE	BOND ENERGY ^a (KCAL/MOL)
Covalent bond	Sharing of electron pairs		50–110
Ionic bond	Attraction of opposite charges		3–7
Hydrogen bond	Sharing of H atom		3–7
Hydrophobic interaction	Interaction of nonpolar substances in the presence of polar substances (especially water)		1–2
van der Waals interaction	Interaction of electrons of nonpolar substances		1

^aBond energy is the amount of energy needed to separate two bonded or interacting atoms under physiological conditions.

sum of the atomic weights of all atoms in the molecule. Looking at the periodic table in Figure 2.2, you can calculate the molecular weights of the three compounds listed above to be 16.04, 18.01, and 342.29, respectively. Molecules that make up living organisms range in molecular weight from two to half a billion, and covalent bonds are common to all.

How are covalent bonds formed in a molecule of methane gas (CH_4)? The carbon atom in this compound has six electrons: two electrons fill its inner shell, and four unpaired electrons travel in its outer shell. Because its outer shell can hold up to eight electrons, carbon can share electrons with up to four other atoms—it can form four covalent bonds (Figure 2.7A). When an atom of carbon reacts with four hydrogen atoms, methane forms. Thanks to electron sharing, the outer shell of methane's carbon atom is now filled with eight electrons, a stable configuration. The outer shell of each of the four hydrogen atoms is also filled. Four covalent bonds—four shared electron pairs—

hold methane together. Figure 2.7B shows several different ways to represent the molecular structure of methane. Table 2.2 shows the covalent bonding capacities of some biologically significant elements.

STRENGTH AND STABILITY Covalent bonds are very strong, meaning that it takes a lot of energy to break them. At temperatures in which life exists, the covalent bonds of biological molecules are quite stable, as are their three-dimensional structures. However, this stability does not preclude change, as we will discover.

ORIENTATION For a given pair of elements—for example, carbon bonded to hydrogen—the length of the covalent bond is always the same. And for a given atom within a molecule, the angle of each covalent bond, with respect to the other bonds, is generally the same. This is true regardless of the type of larger molecule that contains the atom. For example, the four filled orbitals around the carbon atom in methane are always distributed in space so that the bonded hydrogens point to the corners of a regular tetrahedron, with carbon in the center (see Figure 2.7B). Even when carbon is bonded to four atoms other than hydrogen, this three-dimensional orientation is more or less maintained. The orientation of covalent bonds in space gives the molecules their three-dimensional geometry, and the shapes of molecules contribute to their biological functions, as we will see in Section 3.1.

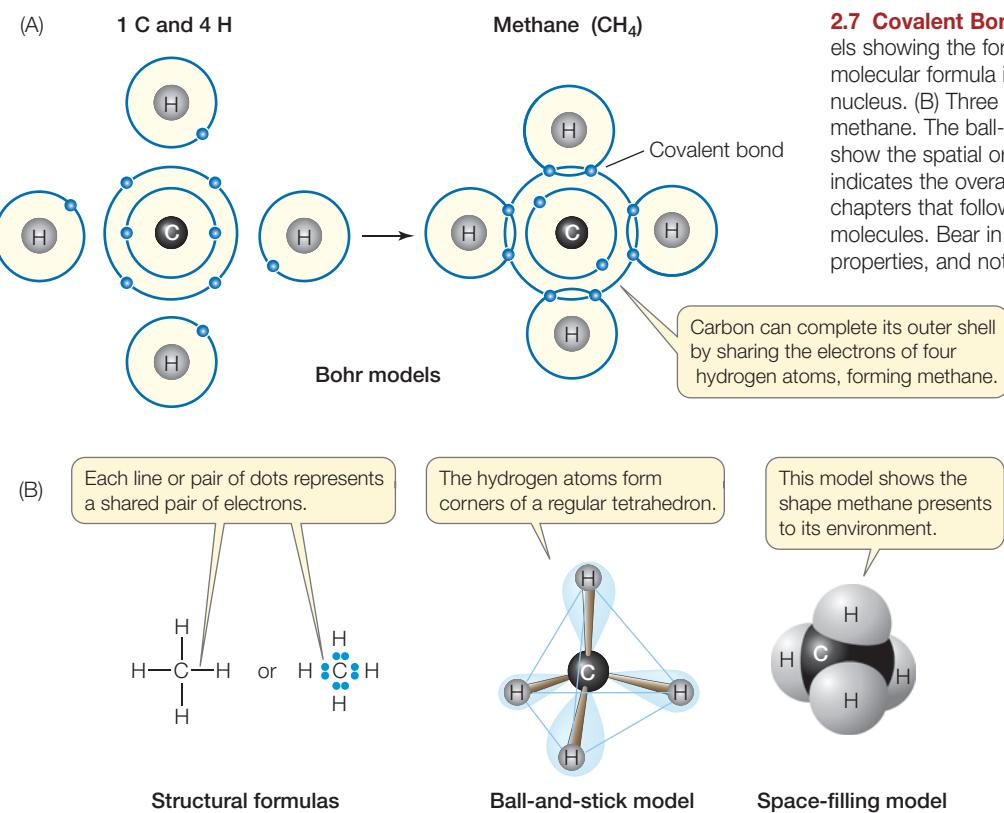
MULTIPLE COVALENT BONDS A covalent bond can be represented by a line between the chemical symbols for the linked atoms:

- A *single bond* involves the sharing of a single pair of electrons (for example, $\text{H}-\text{H}$ or $\text{C}-\text{H}$).

TABLE 2.2

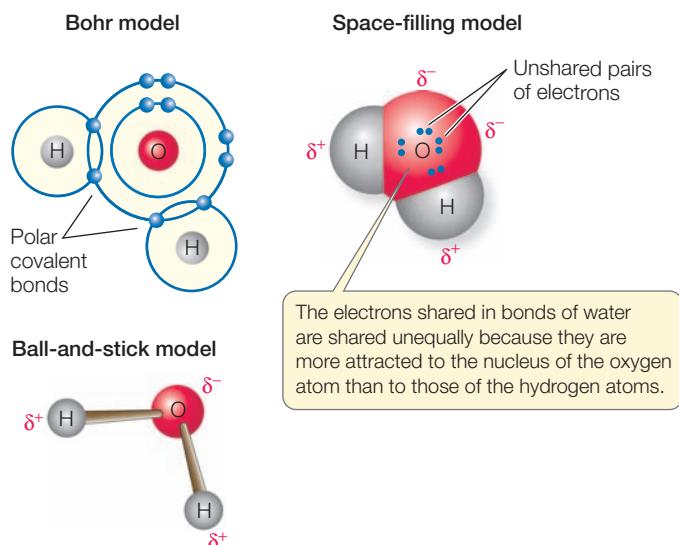
Covalent Bonding Capabilities of Some Biologically Important Elements

ELEMENT	USUAL NUMBER OF COVALENT BONDS
Hydrogen (H)	1
Oxygen (O)	2
Sulfur (S)	2
Nitrogen (N)	3
Carbon (C)	4
Phosphorus (P)	5



- A *double bond* involves the sharing of four electrons (two pairs) ($\text{C}=\text{C}$).
- Triple bonds*—six shared electrons—are rare, but there is one in nitrogen gas ($\text{N}\equiv\text{N}$), which is the major component of the air we breathe.

UNEQUAL SHARING OF ELECTRONS If two atoms of the same element are covalently bonded, there is an equal sharing of the pair(s) of electrons in their outermost shells. However, when the two atoms are of different elements, the sharing is not nec-



2.7 Covalent Bonding Can Form Compounds (A) Bohr models showing the formation of covalent bonds in methane, whose molecular formula is CH_4 . Electrons are shown in shells around the nucleus. (B) Three additional ways of representing the structure of methane. The ball-and-stick model and the space-filling model show the spatial orientations of the bonds. The space-filling model indicates the overall shape and surface of the molecule. In the chapters that follow, different conventions will be used to depict molecules. Bear in mind that these are models to illustrate certain properties, and not the most accurate portrayal of reality.

essarily equal. One nucleus may exert a greater attractive force on the electron pair than the other nucleus, so that the pair tends to be closer to that atom.

The attractive force that an atomic nucleus exerts on electrons in a covalent bond is called its **electronegativity**. The electronegativity of a nucleus depends on how many positive charges it has (nuclei with more protons are more positive and thus more attractive to electrons) and on the distances between the electrons in the bond and the nucleus (the closer the electrons, the greater the electronegative pull). **Table 2.3** shows the electronegativities (which are calculated to produce dimensionless quantities) of some elements important in biological systems.

If two atoms are close to each other in electronegativity, they will share electrons equally in what is called a *nonpolar covalent bond*. Two oxygen atoms, for example, each with an electronegativity of 3.5, will share electrons equally. So will two hydrogen atoms (each with an electronegativity of 2.1). But when hydrogen bonds with oxygen to form water, the electrons involved are unequally shared: they tend to be nearer to the oxygen nucleus because it is the more electronegative of the two. When electrons are drawn to one nucleus more than to the other, the result is a *polar covalent bond* (Figure 2.8).

2.8 Water's Covalent Bonds Are Polar These three representations all illustrate polar covalent bonding in water (H_2O). When atoms with different electronegativities, such as oxygen and hydrogen, form a covalent bond, the electrons are drawn to one nucleus more than to the other. A molecule held together by such a polar covalent bond has partial (δ^+ and δ^-) charges at different surfaces. In water, the shared electrons are displaced toward the oxygen atom's nucleus.

TABLE 2.3
Some Electronegativities

ELEMENT	ELECTRONEGATIVITY
Oxygen (O)	3.5
Chlorine (Cl)	3.1
Nitrogen (N)	3.0
Carbon (C)	2.5
Phosphorus (P)	2.1
Hydrogen (H)	2.1
Sodium (Na)	0.9
Potassium (K)	0.8

Because of this unequal sharing of electrons, the oxygen end of the hydrogen–oxygen bond has a slightly negative charge (symbolized by δ^- and spoken of as “delta negative,” meaning a partial unit of charge), and the hydrogen end has a slightly positive charge (δ^+). The bond is **polar** because these opposite charges are separated at the two ends, or poles, of the bond. The partial charges that result from polar covalent bonds produce polar molecules or polar regions of large molecules. Polar bonds within molecules greatly influence the interactions that they have with other polar molecules. Water (H_2O) is a polar compound, and this polarity has significant effects on its physical properties and chemical reactivity, as we will see in later chapters.

Ionic bonds form by electrical attraction

When one interacting atom is much more electronegative than the other, a complete transfer of one or more electrons may take place. Consider sodium (electronegativity 0.9) and chlorine (3.1). A sodium atom has only one electron in its outermost shell; this condition is unstable. A chlorine atom has seven electrons in its outermost shell—another unstable condition. Since the electronegativity of chlorine is so much greater than that of sodium, any electrons involved in bonding will tend to transfer completely from sodium’s outermost shell to that of chlorine (Figure 2.9). This reaction between sodium and chlorine makes the resulting atoms more stable because they both have eight fully paired electrons in their outer shells. The result is two *ions*.

Ions are electrically charged particles that form when atoms gain or lose one or more electrons:

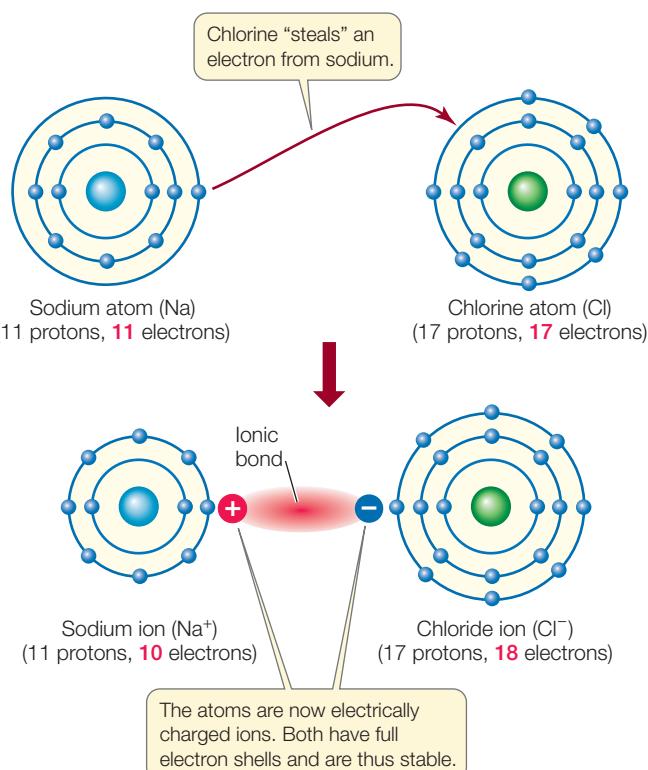
- The sodium ion (Na^+) in our example has a +1 unit of charge because it has one less electron than it has protons. The outermost electron shell of the sodium ion is full, with eight electrons, so the ion is stable. Positively charged ions are called **cations**.
- The chloride ion (Cl^-) has a -1 unit of charge because it has one more electron than it has protons. This additional electron gives Cl^- a stable outermost shell with eight electrons. Negatively charged ions are called **anions**.

Some elements can form ions with multiple charges by losing or gaining *more than one* electron. Examples are Ca^{2+} (the cal-

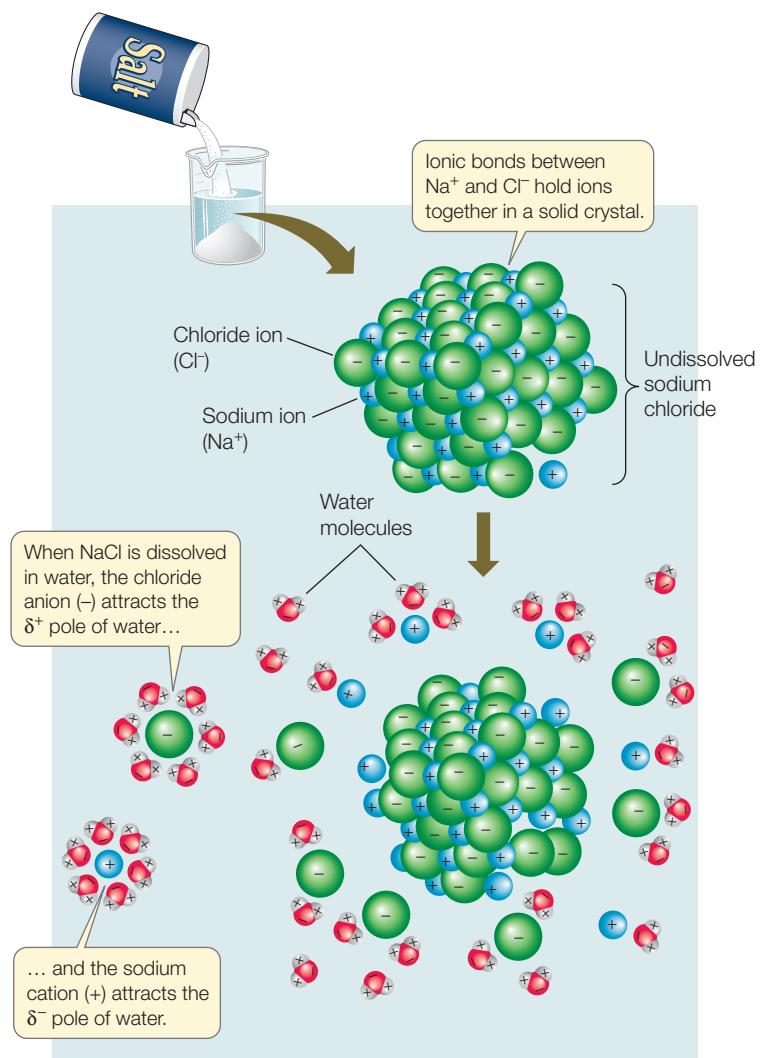
cium ion, a calcium atom that has lost two electrons) and Mg^{2+} (the magnesium ion). Two biologically important elements can each yield more than one stable ion. Iron yields Fe^{2+} (the ferrous ion) and Fe^{3+} (the ferric ion), and copper yields Cu^{+} (the cuprous ion) and Cu^{2+} (the cupric ion). Groups of covalently bonded atoms that carry an electric charge are called *complex ions*; examples include NH_4^+ (the ammonium ion), SO_4^{2-} (the sulfate ion), and PO_4^{3-} (the phosphate ion). Once formed, ions are usually stable and no more electrons are lost or gained.

Ionic bonds are bonds formed as a result of the electrical attraction between ions bearing opposite charges. Ions can form bonds that result in stable solid compounds, which are referred to by the general term *salts*. Examples are sodium chloride ($NaCl$) and potassium phosphate (K_3PO_4). In sodium chloride—familiar to us as table salt—cations and anions are held together by ionic bonds. In solids, the ionic bonds are strong because the ions are close together. However, when ions are dispersed in water, the distance between them can be large; the strength of their attraction is thus greatly reduced. Under the conditions in living cells, an ionic attraction is less strong than a nonpolar covalent bond (see Table 2.1).

Not surprisingly, ions can interact with polar molecules, since they both carry electric charges. Such an interaction results when a solid salt such as $NaCl$ dissolves in water. Water molecules surround the individual ions, separating them (Figure



2.9 Formation of Sodium and Chloride Ions When a sodium atom reacts with a chlorine atom, the more electronegative chlorine fills its outermost shell by “stealing” an electron from the sodium. In so doing, the chlorine atom becomes a negatively charged chloride ion (Cl^-). With one less electron, the sodium atom becomes a positively charged sodium ion (Na^+).



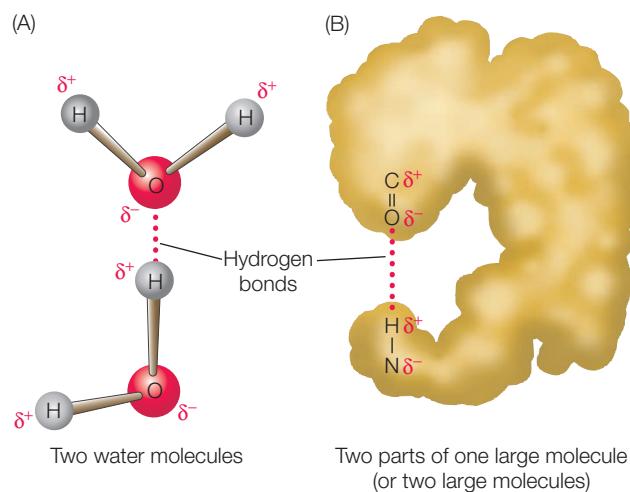
2.10 Water Molecules Surround Ions When an ionic solid dissolves in water, polar water molecules cluster around the cations and anions, preventing them from re-associating.

2.10. The negatively charged chloride ions attract the positive poles of the water molecules, while the positively charged sodium ions attract the negative poles of the water molecules. This is one of the special properties of water molecules, due to their polarity.

Hydrogen bonds may form within or between molecules with polar covalent bonds

In liquid water, the negatively charged oxygen (δ^-) atom of one water molecule is attracted to the positively charged hydrogen (δ^+) atoms of another water molecule (Figure 2.11A). The bond resulting from this attraction is called a **hydrogen bond**. Hydrogen bonds are not restricted to water molecules; they may also form between a strongly electronegative atom and a hydrogen atom that is covalently bonded to a different electronegative atom, as shown in Figure 2.11B.

A hydrogen bond is weaker than most ionic bonds because its formation is due to partial charges (δ^+ and δ^-). It is much weaker than a covalent bond between a hydrogen atom and an oxygen atom (see Table 2.1). Although individual hydrogen bonds are weak, many of them can form within one molecule or



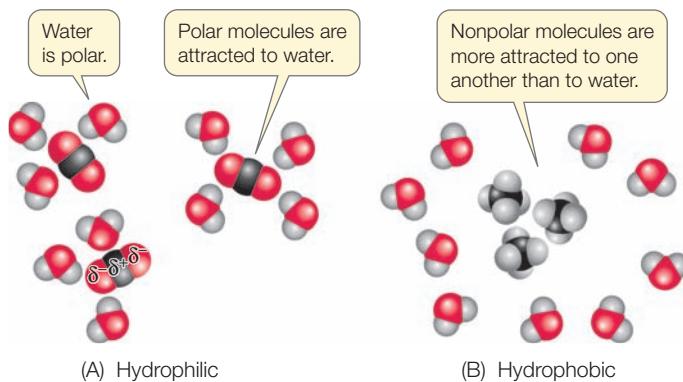
2.11 Hydrogen Bonds Can Form Between or Within Molecules

(A) A hydrogen bond between two molecules is an attraction between a negative charge on one molecule and the positive charge on a hydrogen atom of the second molecule. (B) Hydrogen bonds can form between different parts of the same large molecule.

between two molecules. In these cases, the hydrogen bonds together have considerable strength, and greatly influence the structure and properties of substances. Later in this chapter we'll see how hydrogen bonding between water molecules contributes to many of the properties that make water so significant for living systems. Hydrogen bonds also play important roles in determining and maintaining the three-dimensional shapes of giant molecules such as DNA and proteins (see Section 3.2).

Polar and nonpolar substances: Each interacts best with its own kind

Just as water molecules can interact with one another through hydrogen bonds, any molecule that is polar can interact with other polar molecules through the weak (δ^+ to δ^-) attractions of hydrogen bonds. If a polar molecule interacts with water in this way, it is called **hydrophilic** ("water-loving") (Figure 2.12A).



2.12 Hydrophilic and Hydrophobic (A) Molecules with polar covalent bonds are attracted to polar water (they are hydrophilic). (B) Molecules with nonpolar covalent bonds show greater attraction to one another than to water (they are hydrophobic).

Nonpolar molecules tend to interact with other nonpolar molecules. For example, carbon (electronegativity 2.5) forms nonpolar bonds with hydrogen (electronegativity 2.1), and molecules containing only hydrogen and carbon atoms—called *hydrocarbon molecules*—are nonpolar. In water these molecules tend to aggregate with one another rather than with the polar water molecules. Therefore, nonpolar molecules are known as **hydrophobic** (“water-hating”), and the interactions between them are called *hydrophobic interactions* (Figure 2.12B). Of course, hydrophobic substances do not really “hate” water; they can form weak interactions with it, since the electronegativities of carbon and hydrogen are not exactly the same. But these interactions are far weaker than the hydrogen bonds between the water molecules, so the nonpolar substances tend to aggregate.

The interactions between nonpolar substances are enhanced by **van der Waals forces**, which occur when the atoms of two nonpolar molecules are in close proximity. These brief interactions result from random variations in the electron distribution in one molecule, which create opposite charge distributions in the adjacent molecule. Although a single van der Waals interaction is brief and weak, the sum of many such interactions over the entire span of a large nonpolar molecule can result in substantial attraction. This makes nonpolar molecules stick together in the polar (aqueous) environment inside organisms. We will see this many times, for example in the structure of biological membranes.

2.2 RECAP

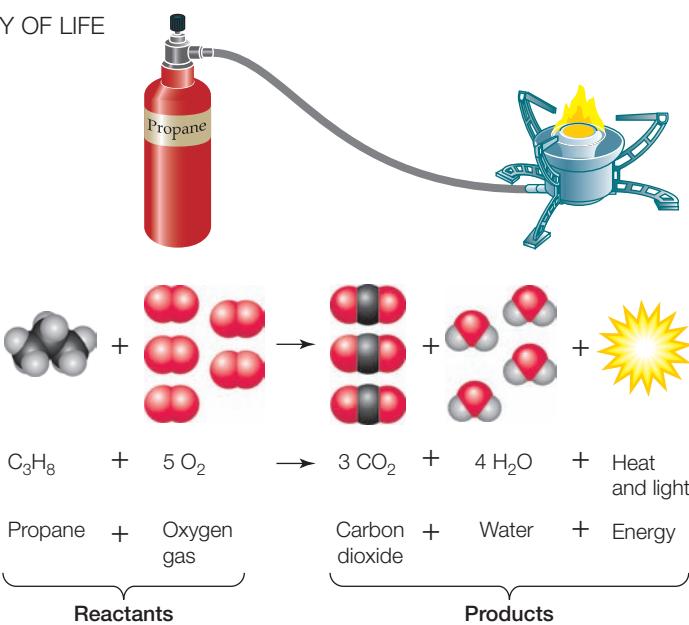
Some atoms form strong covalent bonds with other atoms by sharing one or more pairs of electrons. Unequal sharing of electrons produces polarity. Other atoms become ions by losing or gaining electrons, and they interact with other ions or polar molecules.

- Why is a covalent bond stronger than an ionic bond? **See pp. 26–28 and Table 2.1**
- How do variations in electronegativity result in the unequal sharing of electrons in polar molecules? **See pp. 27–28 and Figure 2.8**
- What is a hydrogen bond and how is it important in biological systems? **See p. 29 and Figure 2.11**

The bonding of atoms into molecules is not necessarily a permanent affair. The dynamic of life involves constant change, even at the molecular level. Let’s look at how molecules interact with one another—how they break up, how they find new partners, and what the consequences of those changes can be.

2.3 How Do Atoms Change Partners in Chemical Reactions?

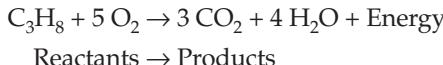
A **chemical reaction** occurs when moving atoms collide with sufficient energy to combine or change their bonding partners. Consider the combustion reaction that takes place in the flame of a propane stove. When propane (C_3H_8) reacts with oxygen gas (O_2), the carbon atoms become bonded to oxygen atoms instead



2.13 Bonding Partners and Energy May Change in a Chemical Reaction

One molecule of propane from this burner reacts with five molecules of oxygen gas to give three molecules of carbon dioxide and four molecules of water. This reaction releases energy in the form of heat and light.

of hydrogen atoms, and the hydrogen atoms become bonded to oxygen instead of carbon (Figure 2.13). As the covalently bonded atoms change partners, the composition of the matter changes; propane and oxygen gas become carbon dioxide and water. This chemical reaction can be represented by the equation



In this equation, the propane and oxygen are the **reactants**, and the carbon dioxide and water are the **products**. In fact, this is a special type of reaction called an oxidation-reduction reaction. Electrons and protons are transferred from propane (the reducing agent) to oxygen (the oxidizing agent) to form water. You will see this kind of reaction involving electron/proton transfer many times in later chapters.

The products of a chemical reaction have very different properties from the reactants. In the case shown in Figure 2.13, the reaction is *complete*: all the propane and oxygen are used up in forming the two products. The arrow symbolizes the direction of the chemical reaction. The numbers preceding the molecular formulas indicate how many molecules are used or produced.

Note that in this and all other chemical reactions, *matter is neither created nor destroyed*. The total number of carbon atoms on the left (3) equals the total number of carbon atoms on the right (3). In other words, the equation is *balanced*. However, there is another aspect of this reaction: the heat and light of the stove’s flame reveal that the reaction between propane and oxygen releases a great deal of energy.

Energy is defined as the capacity to do work, but in the context of chemical reactions, it can be thought of as the capacity for change. Chemical reactions do not create or destroy energy, but *changes in the form of energy* usually accompany chemical reactions.

In the reaction between propane and oxygen, a large amount of heat energy is released. This energy was present in another form, called *potential chemical energy*, in the covalent bonds within

the propane and oxygen gas molecules. Not all reactions release energy; indeed, many chemical reactions require that energy be supplied from the environment. Some of this energy is then stored as potential chemical energy in the bonds formed in the products. We will see in future chapters how reactions that release energy and reactions that require energy can be linked together.

Many chemical reactions take place in living cells, and some of these have a lot in common with the oxidation-reduction reaction that happens in the combustion of propane. In cells, the reactants are different (they may be sugars or fats), and the reactions proceed by many intermediate steps that permit the released energy to be harvested and put to use by the cells. But the products are the same: carbon dioxide and water. We will discuss energy changes, oxidation-reduction reactions, and several other types of chemical reactions that are prevalent in living systems in Part Three of this book.

2.3 RECAP

In a chemical reaction, a set of reactants is converted to a set of products with different chemical compositions. This is accomplished by breaking and making bonds. Reactions may release energy or require its input.

- Explain how a chemical equation is balanced. See p. 30 and Figure 2.13
- How can the form of energy change during a chemical reaction? See p. 30

We will present and discuss energy changes, oxidation-reduction reactions, and several other types of chemical reactions that are prevalent in living systems in Part Two of this book. First, however, we must understand the unique properties of the substance in which most biochemical reactions take place: water.

2.4 What Makes Water So Important for Life?

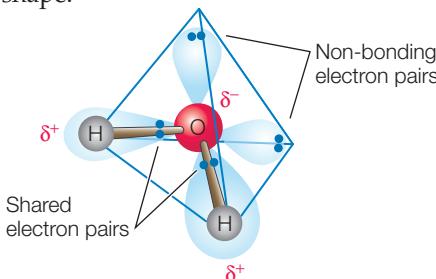
Water is an unusual substance with unusual properties. Under conditions on Earth, water exists in solid, liquid, and gas forms, all of which have relevance to living systems. Water allows chemical reactions to occur inside living organisms, and it is necessary for the formation of certain biological structures. In this section we will explore how the structure and interactions of water molecules make water essential to life.

2.14 Hydrogen Bonding and the Properties of Water

Hydrogen bonding exists between the molecules of water in both its liquid and solid states. Ice is more structured but less dense than liquid water, which is why ice floats. Water forms a gas when its hydrogen bonds are broken and the molecules move farther apart.

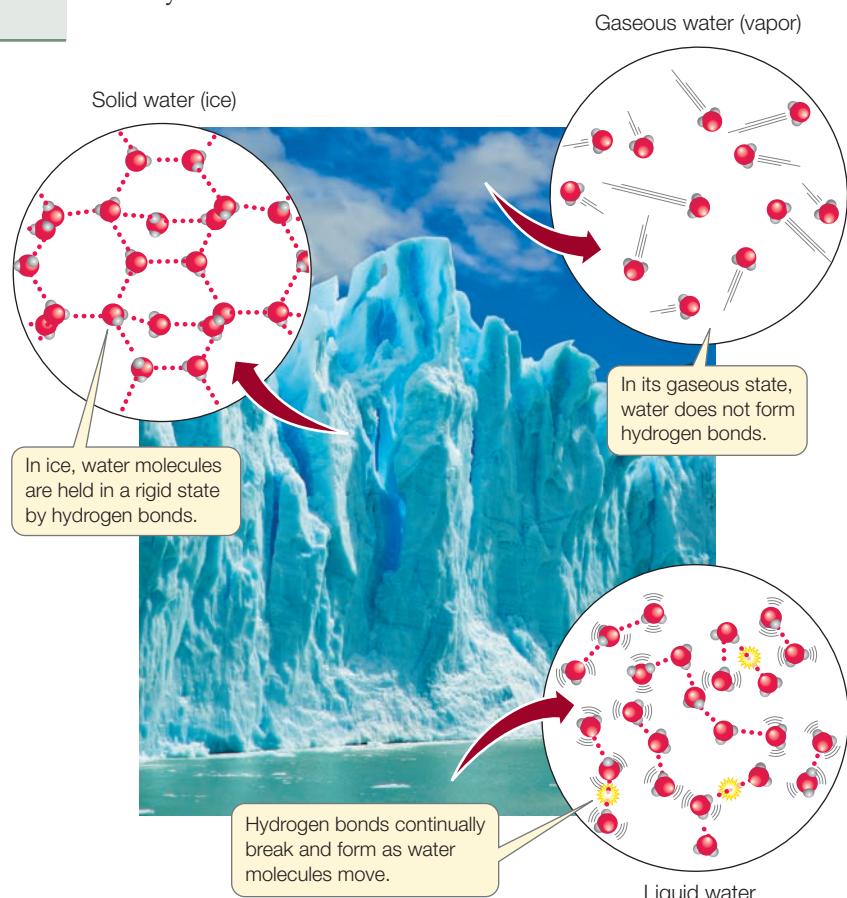
Water has a unique structure and special properties

The molecule H_2O has unique chemical features. As we have already learned, water is a polar molecule that can form hydrogen bonds. The four pairs of electrons in the outer shell of the oxygen atom repel one another, giving the water molecule a tetrahedral shape:



These chemical features explain some of the interesting properties of water, such as the ability of ice to float, the melting and freezing temperatures of water, the ability of water to store heat, the formation of water droplets, and water's ability to dissolve—and not dissolve—many substances.

ICE FLOATS In water's solid state (ice), individual water molecules are held in place by hydrogen bonds. Each molecule is bonded to four other molecules in a rigid, crystalline structure (**Figure 2.14**). Although the molecules are held firmly in place, they are not as tightly packed as they are in liquid water. In other words, *solid water is less dense than liquid water*, which is why ice floats.



Think of the biological consequences if ice were to sink in water. A pond would freeze from the bottom up, becoming a solid block of ice in winter and killing most of the organisms living there. Once the whole pond is frozen, its temperature could drop well below the freezing point of water. But in fact ice floats, forming an insulating layer on the top of the pond, and reducing heat flow to the cold air above. Thus fish, plants, and other organisms in the pond are not subjected to temperatures lower than 0°C, which is the freezing point of pure water.

MELTING, FREEZING, AND HEAT CAPACITY Compared with many other substances that have molecules of similar size, ice requires a great deal of heat energy to melt. This is because so many hydrogen bonds must be broken in order for water to change from solid to liquid. In the opposite process—freezing—a great deal of energy is released to the environment.

This property of water contributes to the surprising constancy of the temperatures found in oceans and other large bodies of water throughout the year. The temperature changes of coastal land masses are also moderated by large bodies of water. Indeed, water helps minimize variations in atmospheric temperature across the planet. This moderating ability is a result of the high *heat capacity* of liquid water, which is in turn a result of its high specific heat.

The **specific heat** of a substance is the amount of heat energy required to raise the temperature of 1 gram of that substance by 1°C. Raising the temperature of liquid water takes a relatively large amount of heat because much of the heat energy is used to break the hydrogen bonds that hold the liquid together. Compared with other small molecules that are liquids, water has a high specific heat.

Water also has a high **heat of vaporization**, which means that a lot of heat is required to change water from its liquid to its gaseous state (the process of *evaporation*). Once again, much of the heat energy is used to break the many hydrogen bonds between the water molecules. This heat must be absorbed from the environment in contact with the water. Evaporation thus has a cooling effect on the environment—whether a leaf, a forest, or an entire land mass. This effect explains why sweating cools the human body: as sweat evaporates from the skin, it uses up some of the adjacent body heat.

COHESION AND SURFACE TENSION In liquid water, individual molecules are able to move about. The hydrogen bonds between the molecules continually form and break (see Figure 2.14). Chemists estimate that this occurs about a trillion times a minute for a single water molecule, making it a truly dynamic structure.

At any given time, a water molecule will form an average of 3.4 hydrogen bonds with other water molecules. These hydrogen bonds explain the *cohesive strength* of liquid water. This cohesive strength, or **cohesion**, is defined as the capacity of water molecules to resist coming apart from one another when placed under tension. Water's cohesive strength permits narrow columns of liquid water to move from the roots to the leaves of tall trees. When water evaporates from the leaves, the entire column moves upward in response to the pull of the molecules at the top.



2.15 Surface Tension Water droplets form “beads” on the surface of a leaf because hydrogen bonds keep the water molecules together. The leaf is coated in a nonpolar wax that does not interact with the water molecules.

The surface of liquid water exposed to the air is difficult to puncture because the water molecules at the surface are hydrogen-bonded to other water molecules below them (Figure 2.15). This *surface tension* of water permits a container to be filled slightly above its rim without overflowing, and it permits insects to walk on the surface of a pond.

Water is an excellent solvent—the medium of life

A human body is over 70 percent water by weight, excluding the minerals contained in bones. Water is the dominant component of virtually all living organisms, and most biochemical reactions take place in this watery, or aqueous, environment.

A **solution** is produced when a substance (the **solute**) is dissolved in a liquid (the **solvent**). If the solvent is water, then the solution is an *aqueous solution*. Many of the important molecules in biological systems are polar, and therefore soluble in water. Many important biochemical reactions occur in aqueous solutions within cells. Biologists study these reactions in order to identify the reactants and products and to determine their amounts:

- *Qualitative analyses* deal with the identification of substances involved in chemical reactions. For example, a qualitative analysis would be used to investigate the steps involved, and the products formed, during the combustion of glucose in living tissues.
- *Quantitative analyses* measure concentrations or amounts of substances. For example, a biochemist would seek to describe *how much* of a certain product is formed during the combustion of a given amount of glucose using a quantitative analysis. What follows is a brief introduction to some of the quantitative chemical terms you will see in this book.

Fundamental to quantitative thinking in chemistry and biology is the concept of the mole. A **mole** is the amount of a substance (in grams) that is numerically equal to its molecular weight.

So a mole of table sugar ($C_{12}H_{22}O_{11}$) weighs about 342 grams; a mole of sodium ion (Na^+) weighs 23 grams; and a mole of hydrogen gas (H_2) weighs 2 grams.

Quantitative analyses do not yield direct counts of molecules. Because the amount of a substance in 1 mole is directly related to its molecular weight, it follows that the number of molecules in 1 mole is constant for all substances. So 1 mole of salt contains the same number of molecules as 1 mole of table sugar. This constant number of molecules in a mole is called **Avogadro's number**, and it is 6.02×10^{23} molecules per mole. Chemists work with moles of substances (which can be weighed out in the laboratory) instead of actual molecules (which are too numerous to be counted). Consider 34.2 grams (just over 1 ounce) of table sugar, $C_{12}H_{22}O_{11}$. This is one-tenth of a mole, or as Avogadro puts it, 6.02×10^{22} molecules.

If you have trouble grasping the concept of a mole, compare it with the concept of a dozen. We buy a dozen eggs or a dozen doughnuts, knowing that we will get 12 of whichever we buy, even though they don't weigh the same or take up the same amount of space.

A chemist can dissolve a mole of sugar (342 g) in water to make 1 liter of solution, knowing that the mole contains 6.02×10^{23} individual sugar molecules. This solution—1 mole of a substance dissolved in water to make 1 liter—is called a 1 molar (1 M) solution. When a physician injects a certain molar concentration of a drug into the bloodstream of a patient, a rough calculation can be made of the actual number of drug molecules that will interact with the patient's cells.

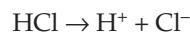
The many molecules dissolved in the water of living tissues are not present at concentrations anywhere near 1 molar. Most are in the micromolar (millionths of a mole per liter of solution; μM) to millimolar (thousandths of a mole per liter; mM) range. Some, such as hormone molecules, are even less concentrated than that. While these molarities seem to indicate very low concentrations, remember that even a 1 μM solution has 6.02×10^{17} molecules of the solute per liter.

Aqueous solutions may be acidic or basic

When some substances dissolve in water, they release *hydrogen ions* (H^+), which are actually single, positively charged protons. Hydrogen ions can attach to other molecules and change their properties. For example, the protons in "acid rain" can damage plants, and you probably have experienced the excess of hydrogen ions that we call "acid indigestion."

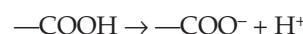
Here we will examine the properties of **acids** (defined as substances that release H^+) and **bases** (defined as substances which accept H^+). We will distinguish between strong and weak acids and bases and provide a quantitative means for stating the concentration of H^+ in solutions: the pH scale.

ACIDS RELEASE H^+ When hydrochloric acid (HCl) is added to water, it dissolves, releasing the ions H^+ and Cl^- :



Because its H^+ concentration has increased, such a solution is *acidic*.

Acids are substances that *release H^+* ions in solution. HCl is an acid, as is H_2SO_4 (sulfuric acid). One molecule of sulfuric acid will ionize to yield two H^+ and one SO_4^{2-} . Biological compounds that contain $-COOH$ (the carboxyl group) are also acids because



Acids that fully ionize in solution, such as HCl and H_2SO_4 are called *strong acids*. However, not all acids ionize fully in water. For example, if acetic acid (CH_3COOH) is added to water, some will dissociate into two ions (CH_3COO^- and H^+), but some of the original acetic acid remains as well. Because the reaction is *not complete*, acetic acid is a *weak acid*.

BASES ACCEPT H^+ Bases are substances that *accept H^+* in solution. Just as with acids, there are strong and weak bases. If NaOH (sodium hydroxide) is added to water, it dissolves and ionizes, releasing OH^- and Na^+ ions:



Because the concentration of OH^- increases and OH^- absorbs H^+ to form water ($OH^- + H^+ \rightarrow H_2O$), such a solution is *basic*. Because this reaction is complete, NaOH is a *strong base*.

Weak bases include the bicarbonate ion (HCO_3^-), which can accept a H^+ ion and become carbonic acid (H_2CO_3), and ammonia (NH_3), which can accept a H^+ and become an ammonium ion (NH_4^+). Biological compounds that contain $-NH_2$ (the amino group) are also bases because



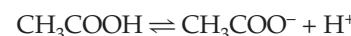
ACID-BASE REACTIONS MAY BE REVERSIBLE When acetic acid is dissolved in water, two reactions happen. First, the acetic acid forms its ions:



Then, once the ions are formed, some of them re-form acetic acid:



This pair of reactions is reversible. A **reversible reaction** can proceed in either direction—left to right or right to left—depending on the relative starting concentrations of the reactants and products. The formula for a reversible reaction can be written using a double arrow:

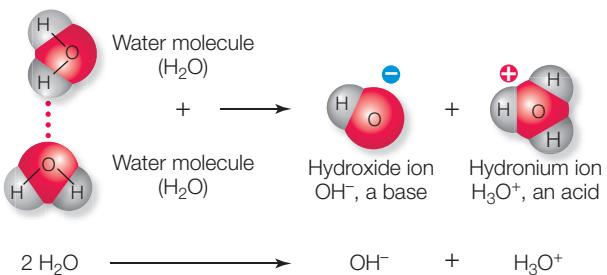


In terms of acids and bases, there are two types of reactions, depending on the extent of the reversibility:

- The ionization of strong acids and bases in water is virtually irreversible.
- The ionization of weak acids and bases in water is somewhat reversible.

WATER IS A WEAK ACID AND A WEAK BASE The water molecule has a slight but significant tendency to ionize into a hydroxide ion (OH^-) and a hydrogen ion (H^+). Actually, two water molecules

participate in this reaction. One of the two molecules “captures” a hydrogen ion from the other, forming a hydroxide ion and a hydronium ion:



The hydronium ion is, in effect, a hydrogen ion bound to a water molecule. For simplicity, biochemists tend to use a modified representation of the ionization of water:



The ionization of water is important to all living creatures. This fact may seem surprising, since only about one water molecule in 500 million is ionized at any given time. But this is less surprising if we focus on the abundance of water in living systems, and the reactive nature of the H^+ ions produced by ionization.

pH: HYDROGEN ION CONCENTRATION Compounds or ions can be acids or bases, and thus, solutions can be acidic or basic. We can measure how acidic or basic a solution is by measuring its concentration of H^+ in moles per liter (its *molarity*; see page 33). Here are some examples:

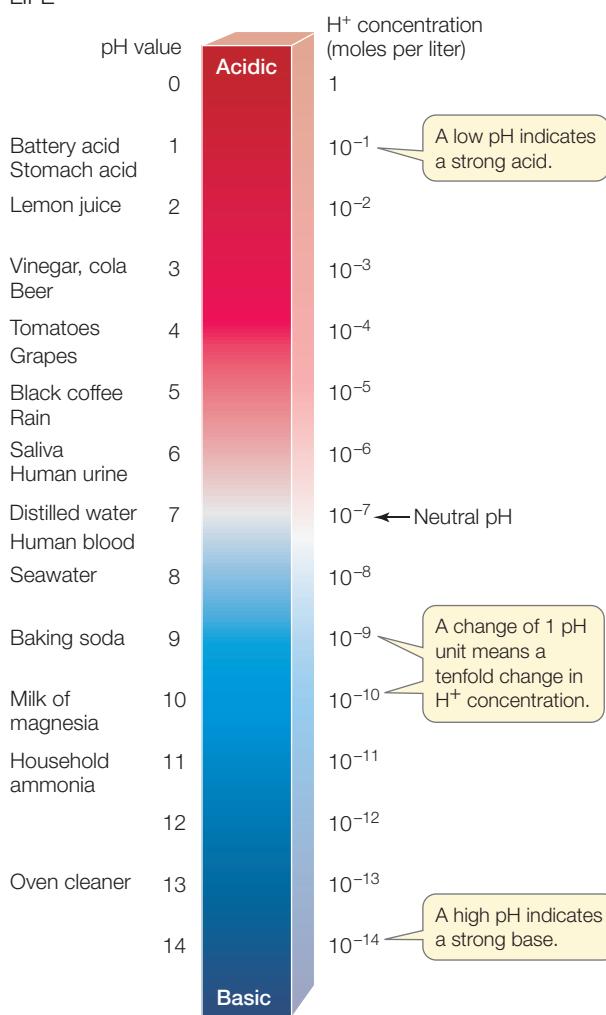
- Pure water has a H^+ concentration of 10^{-7} M .
- A 1 M HCl solution has a H^+ concentration of 1 M (recall that all the HCl dissociates into its ions).
- A 1 M NaOH solution has a H^+ concentration of 10^{-14} M .

This is a very wide range of numbers to work with—think about the decimals! It is easier to work with the *logarithm* of the H^+ concentration, because logarithms compress this range: the \log_{10} of 100, for example is 2, and the \log_{10} of 0.01 is -2. Because most H^+ concentrations in living systems are less than 1M, their \log_{10} values are negative. For convenience, we convert these negative numbers into positive ones, by using the *negative* of the logarithm of the H^+ molar concentration (the molar concentration is designated by square brackets: $[\text{H}^+]$). This number is called the **pH** of the solution.

Since the H^+ concentration of pure water is 10^{-7} M , its pH is $-\log(10^{-7}) = -(−7)$, or 7. A smaller negative logarithm means a larger number. In practical terms, a lower pH means a higher H^+ concentration, or greater acidity. In 1 M HCl, the H^+ concentration is 1 M, so the pH is the negative logarithm of 1 ($-\log 10^0$), or 0. The pH of 1 M NaOH is the negative logarithm of 10^{-14} , or 14.

A solution with a pH of less than 7 is acidic—it contains more H^+ ions than OH^- ions. A solution with a pH of 7 is *neutral* (without net charge), and a solution with a pH value greater than 7 is basic. **Figure 2.16** shows the pH values of some common substances.

Why is this discussion of pH so important in biology? Many biologically important molecules contain charged groups



2.16 pH Values of Some Familiar Substances

(e.g., $-\text{COO}^-$) that can interact with the polar regions of water to form their structures. But these groups can combine with H^+ or other ions in their environment to form uncharged groups (e.g., $-\text{COOH}$, see above). These uncharged groups have much less tendency to interact with water. If such a group is part of a larger molecule, it might now induce the molecule to fold in such a way that it stays away from water because it is hydrophobic. In a more acidic environment, a negatively charged group such as $-\text{COO}^-$ is more likely to combine with H^+ . So the pH of a biological tissue is a key to the three-dimensional structures of many of its constituent molecules. Organisms do all they can to minimize changes in the pH of their watery medium. An important way to do this is with buffers.

BUFFERS The maintenance of internal constancy—*homeostasis*—is a hallmark of all living things and extends to pH. As we mentioned earlier, if biological molecules lose or gain H^+ ions their properties can change, thus upsetting homeostasis. Internal constancy is achieved with buffers: solutions that maintain a relatively constant pH even when substantial amounts of acid or base are added. How does this work?

A **buffer** is a solution of a weak acid and its corresponding base—for example, carbonic acid (H_2CO_3) and bicarbonate ions (HCO_3^-). If an acid is added to a solution containing this buffer,

2.17 Buffers Minimize Changes in pH With increasing amounts of added base, the overall slope of a graph of pH is downward. Without a buffer, the slope is steep. Inside the buffering range of an added buffer, however, the slope is shallow. At very high and very low values of pH, where the buffer is ineffective, the slopes are much steeper.

not all the H⁺ ions from the acid stay in solution. Instead, many of them combine with the bicarbonate ions to produce more carbonic acid:



This reaction uses up some of the H⁺ ions in the solution and decreases the acidifying effect of the added acid. If a base is added, the reaction essentially reverses. Some of the carbonic acid ionizes to produce bicarbonate ions and more H⁺, which counteracts some of the added base. In this way, the buffer minimizes the effect that an added acid or base has on pH. This buffering system is present in the blood, where it is important for preventing significant changes in pH that could disrupt the ability of the blood to carry vital oxygen to tissues. A given amount of acid or base causes a smaller pH change in a buffered solution than in a non-buffered one (Figure 2.17).

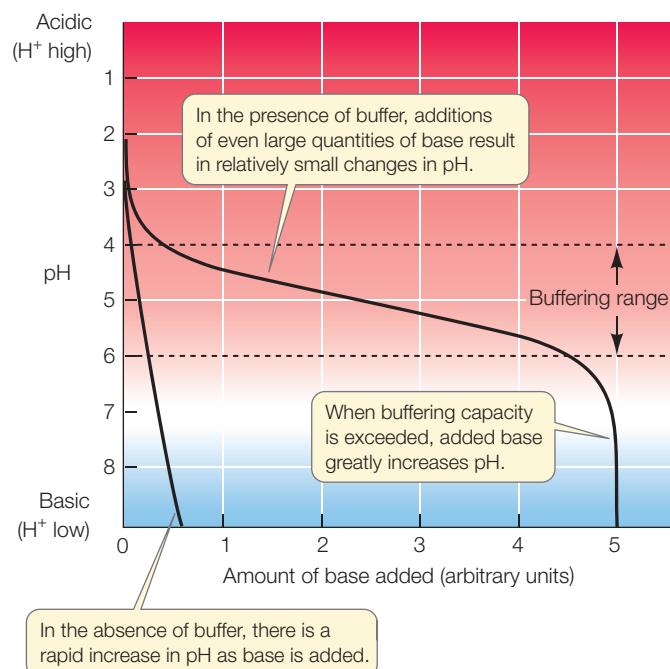
Buffers illustrate an important chemical principle of reversible reactions, called the *law of mass action*. Addition of a reactant on one side of a reversible system drives the reaction in the direction that uses up that compound. In the case of buffers, addition of an acid drives the reaction in one direction; addition of a base drives the reaction in the other direction.

We use a buffer to relieve the common problem of indigestion. The lining of the stomach constantly secretes hydrochloric acid, making the stomach contents acidic. Excessive stomach acid inhibits digestion and causes discomfort. We can relieve this discomfort by ingesting a salt such as NaHCO₃ ("bicarbonate of soda"), which acts as a buffer.

2.4 RECAP

Most of the chemistry of life occurs in water, which has molecular properties that make it suitable for its important biochemical roles. A special property of water is its ability to ionize (release hydrogen ions). The presence of hydrogen ions in solution can change the properties of biological molecules.

- Describe some of the biologically important properties of water arising from its molecular structure. See pp. 31–32 and Figure 2.14
- What is a solution, and why do we call water "the medium of life"? See pp. 32–33
- What is the relationship between hydrogen ions, acids, and bases? Explain what the pH scale measures. See pp. 33–34 and Figure 2.16
- How does a buffer work, and why is buffering important to living systems? See pp. 34–35 and Figure 2.17



An Overview and a Preview

Now that we have covered the major properties of atoms and molecules, let's review them and see how these properties relate to the major molecules of biological systems.

- *Molecules vary in size.* Some are small, such as those of hydrogen gas (H₂) and methane (CH₄). Others are larger, such as a molecule of table sugar (C₁₂H₂₂O₁₁), which has 45 atoms. Still others, especially proteins and nucleic acids, are gigantic, containing tens of thousands or even millions of atoms.
- *All molecules have a specific three-dimensional shape.* For example, the orientations of the bonding orbitals around the carbon atom give the methane molecule (CH₄) the shape of a regular tetrahedron (see Figure 2.7B). Larger molecules have complex shapes that result from the numbers and kinds of atoms present, and the ways in which they are linked together. Some large molecules, such as the protein hemoglobin (the oxygen carrier in red blood cells), have compact, ball-like shapes. Others, such as the protein called keratin that makes up your hair, have long, thin, ropelike structures. Their shapes relate to the roles these molecules play in living cells.
- *Molecules are characterized by certain chemical properties* that determine their biological roles. Chemists use the characteristics of composition, structure (three-dimensional shape), reactivity, and solubility to distinguish a pure sample of one molecule from a sample of a different molecule. The presence of certain groups of atoms can impart distinctive chemical properties to a molecule.

Between the small molecules discussed in this chapter and the world of the living cell are the macromolecules. These larger molecules—proteins, lipids, carbohydrates, and nucleic acids—will be discussed in the next two chapters.

CHAPTER SUMMARY

2.1 How Does Atomic Structure Explain the Properties of Matter?

- Matter is composed of atoms. Each **atom** consists of a positively charged **nucleus** made up of **protons** and **neutrons**, surrounded by **electrons** bearing negative charges. **Review Figure 2.1**
- The number of protons in the nucleus defines an **element**. There are many elements in the universe, but only a few of them make up the bulk of living organisms: C, H, O, P, N, and S. **Review Figure 2.2**
- Isotopes** of an element differ in their numbers of neutrons. **Radioisotopes** are radioactive, emitting radiation as they break down.
- Electrons are distributed in **shells**, which are volumes of space defined by specific numbers of orbitals. Each **orbital** contains a maximum of two electrons. **Review Figures 2.4 and 2.5**, **WEB ACTIVITY 2.1**
- In losing, gaining, or sharing electrons to become more stable, an atom can combine with other atoms to form a **molecule**.

2.2 How Do Atoms Bond to Form Molecules?

SEE ANIMATED TUTORIAL 2.1

- A **chemical bond** is an attractive force that links two atoms together in a molecule. **Review Table 2.1**
- A **compound** is a substance made up of molecules with two or more elements bonded together in a fixed ratio, such as water (H_2O) or table sugar ($C_6H_{12}O_6$).
- Covalent bonds** are strong bonds formed when two atoms share one or more pairs of electrons. **Review Figure 2.6**
- When two atoms of unequal electronegativity bond with each other, a **polar** covalent bond is formed. The two ends, or poles, of the bond have partial charges (δ^+ or δ^-). **Review Figure 2.8**
- Ions** are electrically charged bodies that form when an atom gains or loses one or more electrons in order to form more stable electron configurations. **Anions** and **cations** are negatively and positively charged ions, respectively. Different charges attract, and like charges repel each other.
- Ionic bonds** are electrical attractions between oppositely charged ions. Ionic bonds are strong in solids (salts), but weaken when the ions are separated from one another in solution. **Review Figure 2.9**
- A **hydrogen bond** is a weak electrical attraction that forms between a δ^+ hydrogen atom in one molecule and a δ^- atom in another molecule (or in another part of a large molecule). Hydrogen bonds are abundant in water.

- Nonpolar molecules interact very little with polar molecules, including water. Nonpolar molecules are attracted to one another by very weak bonds called **van der Waals forces**.

2.3 How Do Atoms Change Partners in Chemical Reactions?

- In **chemical reactions**, atoms combine or change their bonding partners. **Reactants** are converted into **products**.
- Some chemical reactions release **energy** as one of their products; other reactions can occur only if energy is provided to the reactants.
- Neither matter nor energy is created or destroyed in a chemical reaction, but both change form. **Review Figure 2.13**
- Some chemical reactions, especially in biology, are reversible. That is, the products formed may be converted back to the reactants.
- In living cells, chemical reactions take place in multiple steps so that the released energy can be harvested for cellular activities.

2.4 What Makes Water So Important for Life?

- Water's molecular structure and its capacity to form hydrogen bonds give it unique properties that are significant for life. **Review Figure 2.14**
- The high **specific heat** of water means that water gains or loses a great deal of heat when it changes state. Water's high **heat of vaporization** ensures effective cooling when water evaporates.
- The **cohesion** of water molecules refers to their capacity to resist coming apart from one another. Hydrogen bonds between water molecules play an essential role in these properties.
- A **solution** is produced when a solid substance (the **solute**) dissolves in a liquid (the **solvent**). Water is the critically important solvent for life.
- Acids** are solutes that release hydrogen ions in aqueous solutions. **Bases** accept hydrogen ions.
- The **pH** of a solution is the negative logarithm of its hydrogen ion concentration. Values lower than pH 7 indicate that a solution is acidic; values above pH 7 indicate a basic solution. **Review Figure 2.16**
- A **buffer** is a mixture of a weak acid and a base that limits changes in the pH of a solution when acids or bases are added.

SELF-QUIZ

- The atomic number of an element
 - equals the number of neutrons in an atom.
 - equals the number of protons in an atom.
 - equals the number of protons minus the number of neutrons.
 - equals the number of neutrons plus the number of protons.
 - depends on the isotope.
- The atomic weight (atomic mass) of an element
 - equals the number of neutrons in an atom.
 - equals the number of protons in an atom.
 - equals the number of electrons in an atom.
 - equals the number of neutrons plus the number of protons.
 - depends on the relative abundances of its electrons and neutrons.
- Which of the following statements about the isotopes of an element is *not* true?
 - They all have the same atomic number.
 - They all have the same number of protons.
 - They all have the same number of neutrons.
 - They all have the same number of electrons.
 - They all have identical chemical properties.

4. Which of the following statements about covalent bonds is *not* true?
 - a. A covalent bond is stronger than a hydrogen bond.
 - b. A covalent bond can form between atoms of the same element.
 - c. Only a single covalent bond can form between two atoms.
 - d. A covalent bond results from the sharing of electrons by two atoms.
 - e. A covalent bond can form between atoms of different elements.
5. Hydrophobic interactions
 - a. are stronger than hydrogen bonds.
 - b. are stronger than covalent bonds.
 - c. can hold two ions together.
 - d. can hold two nonpolar molecules together.
 - e. are responsible for the surface tension of water.
6. Which of the following statements about water is *not* true?
 - a. It releases a large amount of heat when changing from liquid into vapor.
 - b. Its solid form is less dense than its liquid form.
 - c. It is the most effective solvent for polar molecules.
 - d. It is typically the most abundant substance in a living organism.
 - e. It takes part in some important chemical reactions.
7. The reaction $\text{HCl} \rightarrow \text{H}^+ + \text{Cl}^-$ in the human stomach is an example of the
 - a. cleavage of a hydrophobic bond.
 - b. formation of a hydrogen bond.
 - c. elevation of the pH of the stomach.
 - d. formation of ions by dissolving an acid.
 - e. formation of polar covalent bonds.
8. The hydrogen bond between two water molecules arises because water is
 - a. polar.
 - b. nonpolar.
 - c. a liquid.
 - d. small.
 - e. hydrophobic.
9. When table salt (NaCl) is added to water,
 - a. a covalent bond is broken.
 - b. an acidic solution is formed.
 - c. the Na^+ and Cl^- ions are separated.
 - d. the Na^+ ions are attracted to the hydrogen atoms of water.
 - e. water molecules surround the Na^+ (but not Cl^-) ions.
10. The three most abundant elements in a human skin cell are
 - a. calcium, carbon, and oxygen.
 - b. carbon, hydrogen, and oxygen.
 - c. carbon, hydrogen, and sodium.
 - d. carbon, nitrogen, and potassium.
 - e. nitrogen, hydrogen, and argon.

FOR DISCUSSION

1. Using the information in the periodic table (Figure 2.2), draw a Bohr model (see Figures 2.5 and 2.7) of silicon dioxide, showing electrons shared in covalent bonds.
2. Compare a covalent bond between two hydrogen atoms with a hydrogen bond between a hydrogen and an oxygen atom, with regard to the electrons involved, the role of polarity, and the strength of the bond.
3. Write an equation describing the combustion of glucose ($\text{C}_6\text{H}_{12}\text{O}_6$) to produce carbon dioxide and water.
4. The pH of the human stomach is about 2.0, while the pH of the small intestine is about 10.0. What are the hydrogen ion concentrations $[\text{H}^+]$ inside these two organs?

ADDITIONAL INVESTIGATION

Would you expect the elemental composition of Earth's crust to be the same as that of the human body? How could you find out?

3

Proteins, Carbohydrates, and Lipids

Molecular fossils

About 68 million years ago, a *Tyrannosaurus rex*, the fearsome dinosaur of movie stardom, died in what is now Wyoming in the United States. Over time, the giant carcass became buried 60 feet below the surface of what geologists call the Hell Creek Formation. In 2003, a thigh bone from the long-dead beast was found by the famous dinosaur hunter/biologist, John Horner from the Museum of the Rockies. Mary Schweitzer, a molecular paleontologist, was visiting Horner's Montana lab from North Carolina State University. She cut into the bone and found that it contained the remnants of soft tissues (such as bone marrow). This discovery was remark-

able, because up until then scientists had thought that after about a million years, all the soft tissues in bone were replaced with minerals.

Back on the east coast, Lewis Cantley, a biochemist at Harvard University, read about Schweitzer's find in a newspaper and saw the possibility for a unique opportunity: for the first time, a scientist would be able to isolate and study the complex molecules of soft tissues from an extinct organism. He asked Schweitzer to send him a sample, and when he and his colleagues analyzed the dinosaur material, they found fragments of protein molecules.

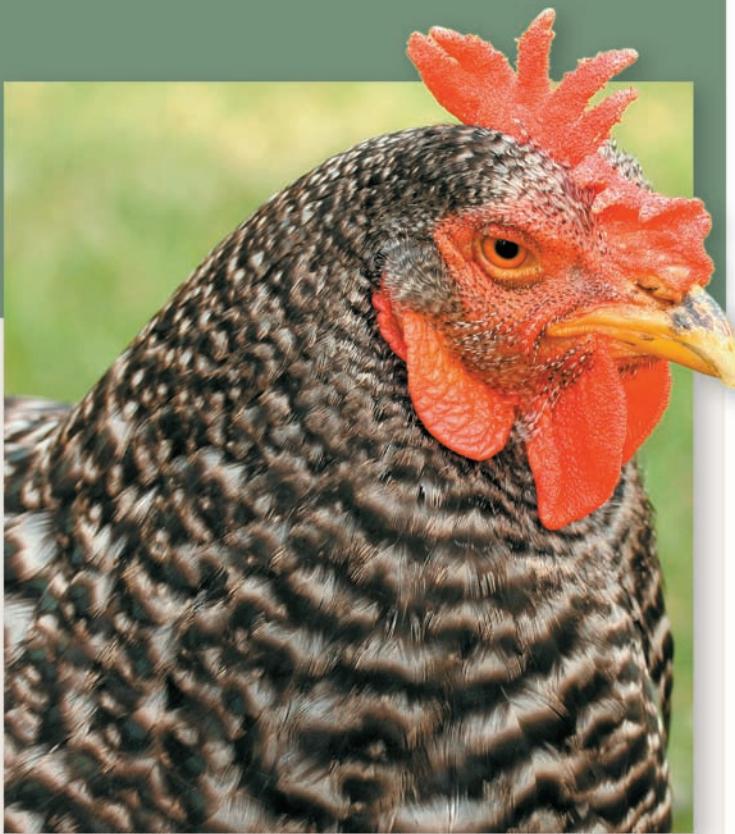
Protein molecules are composed of long chains of individual molecules called amino acids. The protein frag-

ments extracted from the *T. rex* bone were identified as collagen, a substance found in many modern animals. Moreover, the identity and specific order of the amino acids in the dinosaur collagen fragments closely matched that of collagen from chickens, and the dinosaur collagen folded into shapes very similar to those of bird collagen. This similarity to birds is not surprising, because, based on other evidence, scientists believe that birds are evolutionarily closely related to dinosaurs. Cantley's molecular analysis further confirmed this belief.

Proteins are one of the four major kinds of large molecules that characterize living systems. These *macromolecules*, which also include *carbohydrates*, *lipids*, and *nucleic acids*, differ in several significant ways from the small molecules and ions described in Chapter 2. First—no surprise—they are larger; the molecular weights of some



Molecular Clues A thigh bone from a *Tyrannosaurus rex* that died 68 million years ago contained fragments of the protein collagen.



Molecular Evolution The sequence of amino acids in collagen dictates the shape the protein folds into. Collagen's amino acid sequence is similar in *T. rex* and in chickens, indicating that the two species share a common evolutionary ancestor.

nucleic acids reach billions of daltons. Second, these molecules all contain carbon atoms, and so belong to a group of what are known as *organic* chemicals. Third, the atoms of individual macromolecules are held together mostly by covalent bonds, which gives them important structural stability and distinctive three-dimensional geometries. These distinctive shapes are the basis of many of the functions of macromolecules, particularly the proteins.

Finally, carbohydrates, proteins, lipids, and nucleic acids are all unique to the living world. None of these molecular classes occurs in inanimate nature. You aren't likely to find protein in a rock—but if you do, you can be sure it came from a living organism.

IN THIS CHAPTER we will describe the chemical and biological properties of proteins, carbohydrates, and lipids. We will identify the components that make up these larger molecules, describe their assembly and geometries, as well as the roles they play in living organisms.

CHAPTER OUTLINE

- 3.1 What Kinds of Molecules Characterize Living Things?**
- 3.2 What Are the Chemical Structures and Functions of Proteins?**
- 3.3 What Are the Chemical Structures and Functions of Carbohydrates?**
- 3.4 What Are the Chemical Structures and Functions of Lipids?**

3.1 What Kinds of Molecules Characterize Living Things?

Four kinds of molecules are characteristic of living things: proteins, carbohydrates, lipids, and nucleic acids. With the exception of the lipids, these *biological molecules* are **polymers** (*poly*, “many”; *mer*, “unit”) constructed by the covalent bonding of smaller molecules called **monomers**. The monomers that make up each kind of biological molecule have similar chemical structures:

- *Proteins* are formed from different combinations of 20 *amino acids*, all of which share chemical similarities.
- *Carbohydrates* can form giant molecules by linking together chemically similar sugar monomers (*monosaccharides*) to form polysaccharides.
- *Nucleic acids* are formed from four kinds of nucleotide monomers linked together in long chains.
- *Lipids* also form large structures from a limited set of smaller molecules, but in this case noncovalent forces maintain the interactions between the lipid monomers.

Polymers with molecular weights exceeding 1,000 grams per mole are considered to be **macromolecules**. The proteins, carbohydrates, and nucleic acids of living systems certainly fall into this category. Although large lipid structures are not polymers in the strictest sense, it is convenient to treat them as a special type of macromolecule (see Section 3.4).

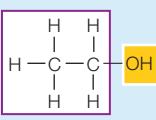
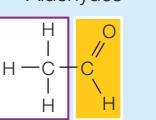
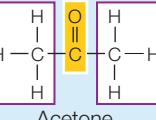
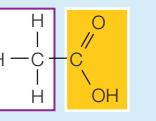
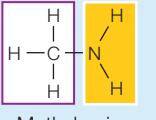
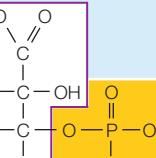
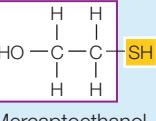
How the macromolecules function and interact with other molecules depends on the properties of certain chemical groups in their monomers, the *functional groups*.

[yourBioPortal.com](#)

[GO TO](#) Animated Tutorial 3.1 • Macromolecules

Functional groups give specific properties to biological molecules

Certain small groups of atoms, called **functional groups**, are consistently found together in very different biological molecules. You will encounter several functional groups repeatedly in your study of biology (Figure 3.1). Each functional group has specific chemical properties and, when it is attached to a larger molecule, it confers those properties on the larger molecule. One of these properties is polarity. Looking at the structures in Figure 3.1, can you determine which functional groups are the most

Functional group	Class of compounds and an example	Properties
Hydroxyl	Alcohols 	Polar. Hydrogen bonds with water to help dissolve molecules. Enables linkage to other molecules by dehydration.
Aldehyde	Aldehydes 	C=O group is very reactive. Important in building molecules and in energy-releasing reactions.
Keto	Ketones 	C=O group is important in carbohydrates and in energy reactions.
Carboxyl	Carboxylic acids 	Acidic. Ionizes in living tissues to form COO^- and H^+ . Enters into dehydration synthesis by giving up OH . Some carboxylic acids important in energy-releasing reactions.
Amino	Amines 	Basic. Accepts H^+ in living tissues to form NH_3^+ . Enters into dehydration synthesis by giving up H^+ .
Phosphate	Organic phosphates 	Negatively charged. Enters into dehydration synthesis by giving up OH . When bonded to another phosphate, hydrolysis releases much energy.
Sulfhydryl	Thiols 	By giving up H, two SH groups can react to form a disulfide bridge, thus stabilizing protein structure.

polar? (Hint: Look for C—O, N—H, and P—O bonds.) The consistent chemical behavior of functional groups helps us understand the properties of the molecules that contain them.

Because macromolecules are so large, they contain many different functional groups (see Figure 3.1). A single large protein may contain hydrophobic, polar, and charged functional groups, each of which gives different specific properties to local sites on the macromolecule. As we will see, sometimes these

3.1 Some Functional Groups Important to Living Systems

Highlighted here are the seven functional groups most commonly found in biologically important molecules. “R” is a variable chemical grouping.

[yourBioPortal.com](#)

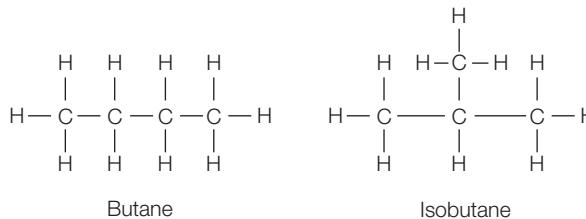
GO TO Web Activity 3.1 • Functional Groups

different groups interact on the same macromolecule. These diverse groups and their properties help determine the shapes of macromolecules as well as how they interact with other macromolecules and with smaller molecules.

Isomers have different arrangements of the same atoms

Isomers are molecules that have the same chemical formula—the same kinds and numbers of atoms—but the atoms are arranged differently. (The prefix *iso-*, meaning “same,” is encountered in many biological terms.) Of the different kinds of isomers, we will consider two: structural isomers and optical isomers.

Structural isomers differ in how their atoms are joined together. Consider two simple molecules, each composed of four carbon and ten hydrogen atoms bonded covalently, both with the formula C_4H_{10} . These atoms can be linked in two different ways, resulting in different molecules:

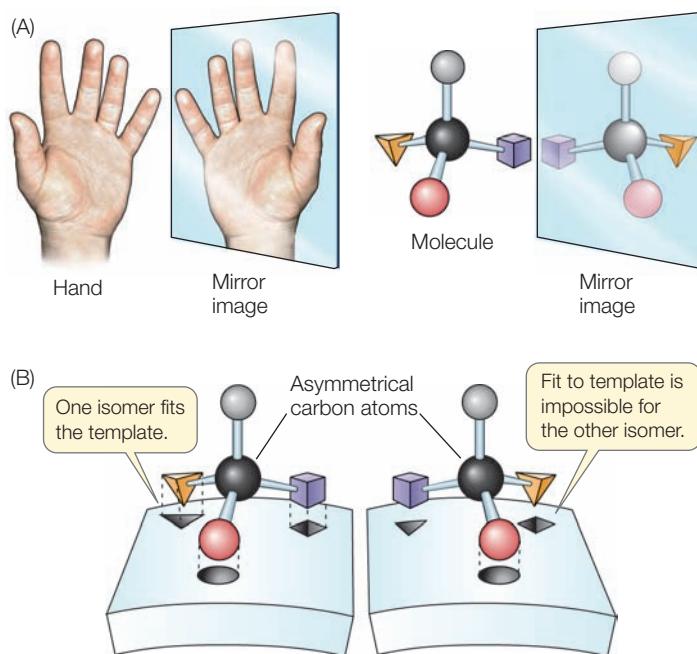


The different bonding relationships in butane and isobutane are distinguished by their structural formulas, and the two molecules have different chemical properties.

Optical isomers occur when a carbon atom has four different atoms or groups of atoms attached to it. This pattern allows two different ways of making the attachments, each the mirror image of the other (Figure 3.2). Such a carbon atom is called an *asymmetrical carbon*, and the two resulting molecules are optical isomers of each other. You can envision your right and left hands as optical isomers. Just as a glove is specific for a particular hand, some biochemical molecules that can interact with one optical isomer of a carbon compound are unable to “fit” the other.

The structures of macromolecules reflect their functions

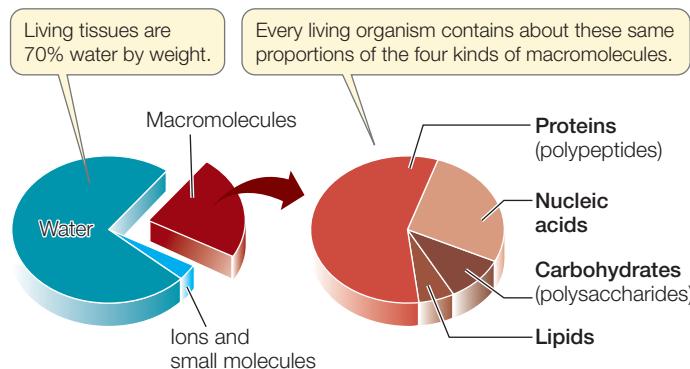
The four kinds of biological macromolecules are present in roughly the same proportions in all living organisms (Figure 3.3). Furthermore, a protein that has a certain function in an apple tree probably has a similar function in a human being because its chemistry is the same wherever it is found. Such *biochemical unity* reflects the evolution of all life from a common ancestor, by descent with modification. An important advantage of biochemical unity is that some organisms can acquire



3.2 Optical Isomers (A) Optical isomers are mirror images of each other. (B) Molecular optical isomers result when four different atoms or groups are attached to a single carbon atom. If a template (representing a larger biological molecule in a living system) is laid out to match the groups on one carbon atom, the groups on that carbon's optical isomer cannot be rotated to fit the same template. This is a source of specificity in biological structure and biochemical transformations.

needed raw materials by eating other organisms. When you eat an apple, the molecules you take in include carbohydrates, lipids, and proteins that can be broken down and rebuilt into the varieties of those molecules needed by humans.

Each type of macromolecule performs some combination of functions, such as energy storage, structural support, protection, catalysis (speeding up a chemical reaction), transport, defense, regulation, movement, and information storage. These roles are not necessarily exclusive; for example, both carbohydrates and proteins can play structural roles, supporting and protecting tissues and organs. However, only the nucleic acids specialize in



3.3 Substances Found in Living Tissues The substances shown here make up the nonmineral components of living tissues (bone would be an example of a mineral component).

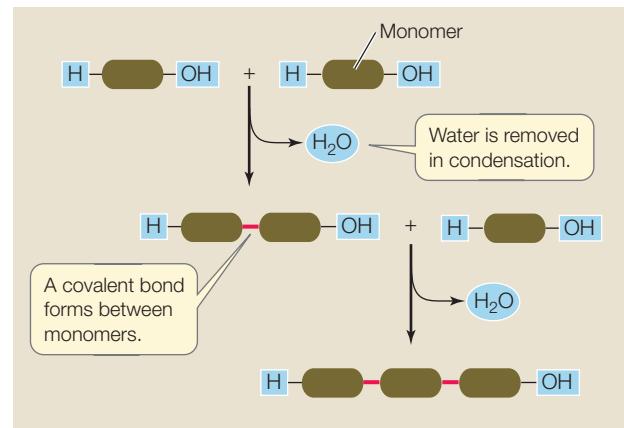
information storage and transmission. These macromolecules function as hereditary material, carrying the traits of both species and individuals from generation to generation.

The functions of macromolecules are directly related to their three-dimensional shapes and to the sequences and chemical properties of their monomers. Some macromolecules fold into compact spherical forms with surface features that make them water-soluble and capable of intimate interaction with other molecules. Some proteins and carbohydrates form long, fibrous systems (such as those found in hair) that provide strength and rigidity to cells and tissues. The long, thin assemblies of proteins such as those in muscles can contract, resulting in movement.

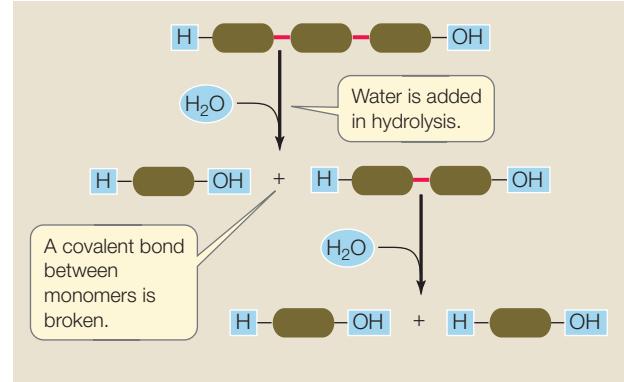
Most macromolecules are formed by condensation and broken down by hydrolysis

Polymers are constructed from monomers by a series of reactions called **condensation reactions** (sometimes called *dehydration* reactions; both terms refer to the loss of water). Condensation reactions result in covalent bonds between monomers. A molecule of water is released with each covalent bond formed (Figure 3.4A). The condensation reactions that produce the different kinds of polymers differ in detail, but in

(A) Condensation



(B) Hydrolysis



3.4 Condensation and Hydrolysis of Polymers (A) Condensation reactions link monomers into polymers and produce water. (B) Hydrolysis reactions break polymers into individual monomers and consume water.

all cases, polymers form only if water molecules are removed and energy is added to the system. In living systems, specific energy-rich molecules supply the necessary energy.

The reverse of a condensation reaction is a **hydrolysis reaction** (*hydro*, “water”; *lysis*, “break”). Hydrolysis reactions result in the breakdown of polymers into their component monomers. Water reacts with the covalent bonds that link the polymer together. For each covalent bond that is broken, a water molecule splits into two ions (H^+ and OH^-), which each become part of one of the products (Figure 3.4B). The linkages between monomers can thus be formed and broken inside living tissues.

3.1 RECAP

The four kinds of large molecules that distinguish living tissues are proteins, lipids, carbohydrates, and nucleic acids. These biological molecules carry out a wide range of life-sustaining functions. Most of them are polymers, made up of linked monomeric subunits. Very large polymers are called macromolecules.

- How do functional groups affect the structure and function of macromolecules? (Keep this question in mind as you read the rest of this chapter.) See pp. 39–40 and Figure 3.1
- Why is biochemical unity, as seen in the proportions of the four types of macromolecules present in all organisms, important for life? See p. 40 and Figure 3.3
- How do monomers link up to make polymers and how do they break down into monomers again? See pp. 41–42 and Figure 3.4

The four types of macromolecules can be seen as the building blocks of life. The unique properties of the nucleic acids will be covered in Chapter 4. The remainder of this chapter describes the structures and functions of the proteins, carbohydrates, and lipids.

3.2 What Are the Chemical Structures and Functions of Proteins?

While all of the kinds of large molecules are essential to the function of organisms, few have such diverse roles as the proteins. In virtually every chapter of this book, you will be studying examples of their extensive functions:

- *Enzymes* are catalytic proteins that speed up biochemical reactions.
- *Defensive proteins* such as antibodies recognize and respond to non-self substances that invade the organism from the environment.
- *Hormonal and regulatory proteins* such as insulin control physiological processes.
- *Receptor proteins* receive and respond to molecular signals from inside and outside the organism.

- *Storage proteins* store chemical building blocks—amino acids—for later use.
- *Structural proteins* such as collagen provide physical stability and movement.
- *Transport proteins* such as hemoglobin carry substances within the organism.
- *Genetic regulatory proteins* regulate when, how, and to what extent a gene is expressed.

Among the functions of macromolecules listed earlier, only two—energy storage and information storage—are not usually performed by proteins.

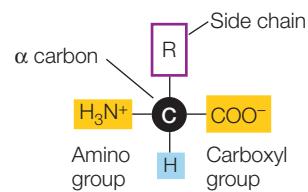
All **proteins** are polymers made up of different proportions and sequences of 20 amino acids. Proteins range in size from small ones such as insulin, which has a molecular weight of 5,733 daltons and 51 amino acids, to huge molecules such as the muscle protein titin, with a molecular weight of 2,993,451 daltons and 26,926 amino acids. All proteins consist of one or more *polypeptide chains*—unbranched (linear) polymer of covalently linked amino acids. The *composition* of a protein refers to the relative amounts of the different amino acids present in its polypeptide chains. Variation in the *sequence* of the amino acids in polypeptide chains is the source of the diversity in protein structure and function, because each chain folds into specific three-dimensional shape that is defined by the precise sequence of the amino acids present in the chain.

Many proteins are made up of more than one polypeptide chain. For example, the oxygen-carrying protein hemoglobin has four chains that are folded separately and come together to make up the functional protein. Proteins can also associate with one another, forming multi-protein complexes that carry out intricate tasks such as DNA synthesis.

To understand the many functions of proteins, we must first explore protein structure. We begin by examining the properties of amino acids and how they link together to form polypeptide chains. Then we will describe how a linear chain of amino acids is consistently folded into a specific, compact, three-dimensional shape. Finally, we will see how this three-dimensional structure provides a definitive physical and chemical environment that influences how other molecules can interact with the protein.

Amino acids are the building blocks of proteins

The amino acids have both a carboxyl functional group and an amino functional group (see Figure 3.1) attached to the same carbon atom, called the α (alpha) carbon. Also attached to the α carbon atom are a hydrogen atom and a **side chain**, or **R group**, designated by the letter R.



The α carbon is asymmetrical because it is bonded to four different atoms or groups of atoms. Therefore, amino acids exist

in two isomeric forms, called D-amino acids and L-amino acids. D and L are abbreviations of the Latin terms for right (*dextro*) and left (*levo*). Only L-amino acids are commonly found in proteins in most organisms, and their presence is an important chemical “signature” of life.

At the pH values commonly found in cells, both the carboxyl and amino groups of amino acids are ionized: the carboxyl group has lost a hydrogen ion, and the amino group has gained one. Thus *amino acids are simultaneously acids and bases*.

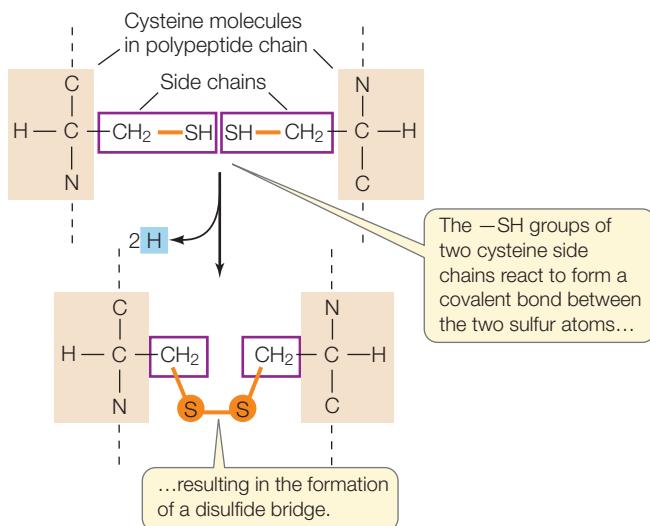
The side chains of amino acids contain functional groups that are important in determining the three-dimensional structure and thus the function of the protein. As **Table 3.1** shows, the 20 amino acids found in living organisms are grouped and distinguished by their side chains:

- The five amino acids that have electrically charged side chains (+1, -1) attract water (are hydrophilic) and attract oppositely charged ions of all sorts.
- The five amino acids that have polar side chains (δ^+ , δ^-) tend to form hydrogen bonds with water and with other polar or charged substances. These amino acids are also hydrophilic.
- Seven amino acids have side chains that are nonpolar hydrocarbons or very slightly modified hydrocarbons. In the watery environment of the cell, these hydrophobic side chains may cluster together in the interior of the protein. These amino acids are hydrophobic.

TABLE 3.1

The Twenty Amino Acids

<p>Amino acids have both three-letter and single-letter abbreviations.</p> <p>A. Amino acids with electrically charged hydrophilic side chains</p> <table border="0"> <thead> <tr> <th colspan="3">Positive +</th><th colspan="3">Negative -</th></tr> </thead> <tbody> <tr> <td>Arginine (Arg; R)</td><td>Histidine (His; H)</td><td>Lysine (Lys; K)</td><td>Aspartic acid (Asp; D)</td><td>Glutamic acid (Glu; E)</td><td></td></tr> <tr> <td> </td><td> </td><td> </td><td> </td><td> </td><td></td></tr> <tr> <td colspan="6"> <p>The general structure of all amino acids is the same... ...but each has a different side chain.</p> </td></tr> </tbody> </table>						Positive +			Negative -			Arginine (Arg; R)	Histidine (His; H)	Lysine (Lys; K)	Aspartic acid (Asp; D)	Glutamic acid (Glu; E)								<p>The general structure of all amino acids is the same... ...but each has a different side chain.</p>					
Positive +			Negative -																										
Arginine (Arg; R)	Histidine (His; H)	Lysine (Lys; K)	Aspartic acid (Asp; D)	Glutamic acid (Glu; E)																									
<p>The general structure of all amino acids is the same... ...but each has a different side chain.</p>																													
<p>B. Amino acids with polar but uncharged side chains (hydrophilic)</p> <table border="0"> <tbody> <tr> <td>Serine (Ser; S)</td> <td>Threonine (Thr; T)</td> <td>Asparagine (Asn; N)</td> <td>Glutamine (Gln; Q)</td> <td>Tyrosine (Tyr; Y)</td> <td></td> </tr> <tr> <td> </td> <td> </td> <td> </td> <td> </td> <td> </td> <td></td> </tr> </tbody> </table>						Serine (Ser; S)	Threonine (Thr; T)	Asparagine (Asn; N)	Glutamine (Gln; Q)	Tyrosine (Tyr; Y)																			
Serine (Ser; S)	Threonine (Thr; T)	Asparagine (Asn; N)	Glutamine (Gln; Q)	Tyrosine (Tyr; Y)																									
<p>C. Special cases</p> <table border="0"> <tbody> <tr> <td>Cysteine (Cys; C)</td> <td>Glycine (Gly; G)</td> <td>Proline (Pro; P)</td> </tr> <tr> <td> </td> <td> </td> <td> </td> </tr> </tbody> </table>						Cysteine (Cys; C)	Glycine (Gly; G)	Proline (Pro; P)																					
Cysteine (Cys; C)	Glycine (Gly; G)	Proline (Pro; P)																											
<p>D. Amino acids with nonpolar hydrophobic side chains</p> <table border="0"> <tbody> <tr> <td>Alanine (Ala; A)</td> <td>Isoleucine (Ile; I)</td> <td>Leucine (Leu; L)</td> <td>Methionine (Met; M)</td> <td>Phenylalanine (Phe; F)</td> <td>Tryptophan (Trp; W)</td> <td>Valine (Val; V)</td> </tr> <tr> <td> </td> </tr> </tbody> </table>						Alanine (Ala; A)	Isoleucine (Ile; I)	Leucine (Leu; L)	Methionine (Met; M)	Phenylalanine (Phe; F)	Tryptophan (Trp; W)	Valine (Val; V)																	
Alanine (Ala; A)	Isoleucine (Ile; I)	Leucine (Leu; L)	Methionine (Met; M)	Phenylalanine (Phe; F)	Tryptophan (Trp; W)	Valine (Val; V)																							



3.5 A Disulfide Bridge Two cysteine molecules in a polypeptide chain can form a disulfide bridge ($-S-S-$) by oxidation (removal of H atoms).

Three amino acids—cysteine, glycine, and proline—are special cases, although the side chains of the latter two are generally hydrophobic.

- The *cysteine* side chain, which has a terminal $-SH$ group, can react with another cysteine side chain in an oxidation reaction to form a covalent bond called a **disulfide bridge**, or *disulfide bond* ($-S-S-$) (Figure 3.5). Disulfide bridges help determine how a polypeptide chain folds.
- The *glycine* side chain consists of a single hydrogen atom and is small enough to fit into tight corners in the interior of a protein molecule, where a larger side chain could not fit.
- Proline* possesses a modified amino group that lacks a hydrogen and instead forms a covalent bond with the hydrocarbon side chain, resulting in a ring structure. This limits both its hydrogen-bonding ability and its ability to rotate about the α carbon. Thus proline is often found where a protein bends or loops.

yourBioPortal.com

GO TO Web Activity 3.2 • Features of Amino Acids

Peptide linkages form the backbone of a protein

When amino acids polymerize, the carboxyl and amino groups attached to the α carbon are the reactive groups. The carboxyl group of one amino acid reacts with the amino group of another, undergoing a condensation reaction that forms a **peptide linkage** (also called a *peptide bond*). Figure 3.6 gives a simplified description of this reaction.

Just as a sentence begins with a capital letter and ends with a period, polypeptide chains have a beginning and an end. The “capital letter” marking the beginning of a polypeptide is the amino group of the first amino acid added to the chain and is known as the *N terminus*. The “period” is the carboxyl group of the last amino acid added; this is the *C terminus*.

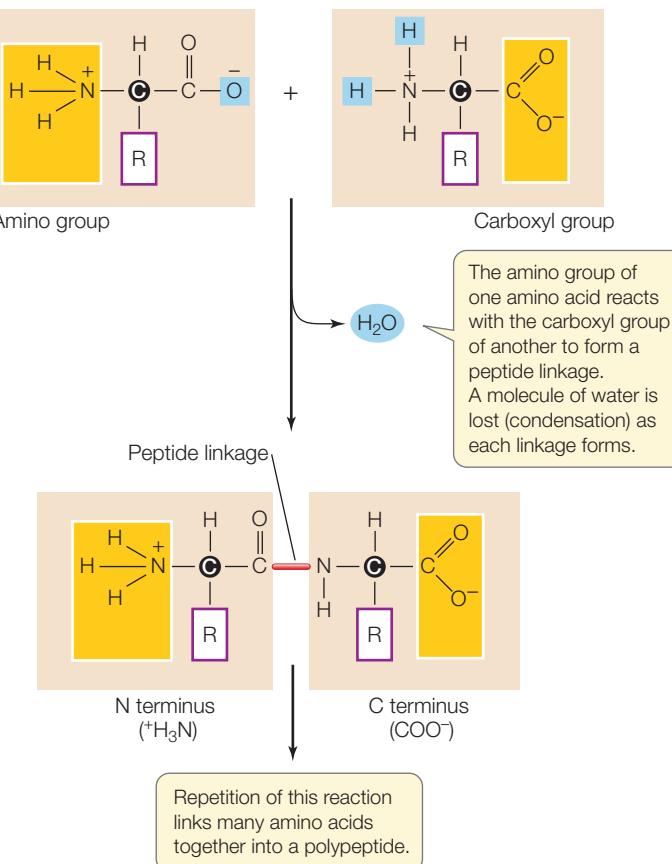
Two characteristics of the peptide bond are especially important in the three-dimensional structure of proteins:

- In the C—N linkage, the adjacent α carbons ($\alpha C-C-N-\alpha C$) are not free to rotate fully, which limits the folding of the polypeptide chain.
- The oxygen bound to the carbon ($C=O$) in the carboxyl group carries a slight negative charge (δ^-), whereas the hydrogen bound to the nitrogen ($N-H$) in the amino group is slightly positive (δ^+). This asymmetry of charge favors hydrogen bonding within the protein molecule itself and with other molecules, contributing to both the structure and the function of many proteins.

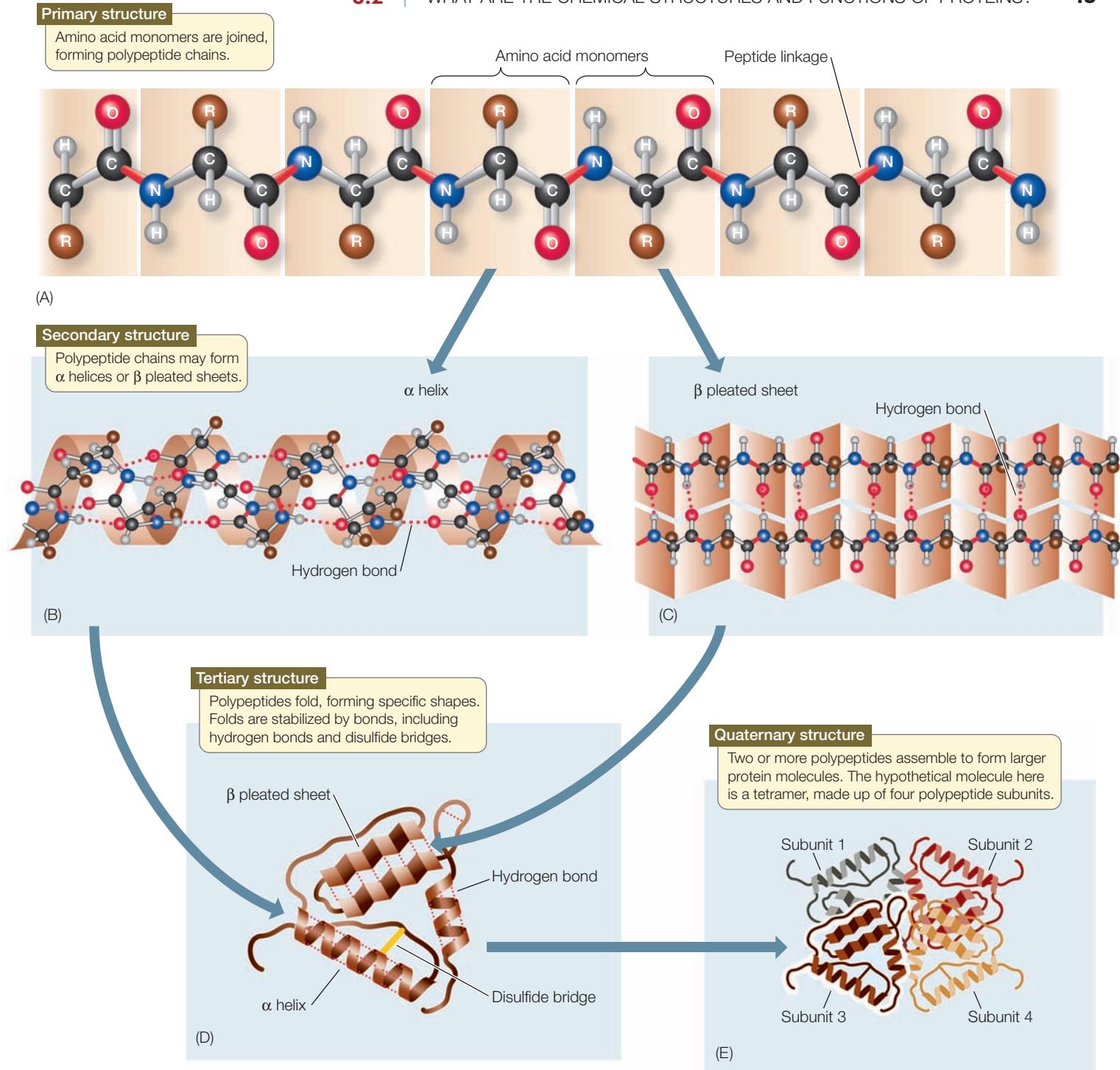
Before we explore the significance of these characteristics of the peptide linkage, however, we will describe the significance of the sequence of amino acids in determining a protein’s structure.

The primary structure of a protein is its amino acid sequence

There are four levels of protein structure: primary, secondary, tertiary, and quaternary. We will consider each of these in turn over the next few pages. The precise sequence of amino acids in a polypeptide chain held together by peptide linkages constitutes the **primary structure** of a protein (Figure 3.7A). The peptide backbone of the polypeptide chain consists of the repeating sequence $-N-C-C-$ made up of the N atom from the



3.6 Formation of Peptide Linkages In living things, the reaction leading to a peptide linkage (also called a peptide bond) has many intermediate steps, but the reactants and products are the same as those shown in this simplified diagram.



amino group, the α carbon atom, and the C atom from the carboxyl group of each amino acid.

Scientists have determined the primary structure of many proteins. The single-letter abbreviations for amino acids (see Table 3.1) are used to record the amino acid sequence of a protein. Here, for example, are the first 20 amino acids (out of a total of 124) in the protein ribonuclease from a cow:

KETAAAKFERQHMDSTSAA

The theoretical number of different proteins is enormous. Since there are 20 different amino acids, there could be $20 \times 20 = 400$ distinct dipeptides (two linked amino acids), and $20 \times 20 \times 20 =$

3.7 The Four Levels of Protein Structure Secondary, tertiary, and quaternary structure all arise from the primary structure of the protein.

8,000 different tripeptides (three linked amino acids). Imagine this process of multiplying by 20 extended to a protein made up of 100 amino acids (which would be considered a small protein). There could be 20^{100} (that's approximately 10^{130}) such small proteins, each with its own distinctive primary structure. How large is the number 20^{100} ? Physicists tell us that there aren't that many electrons in the entire universe.

At the higher levels of protein structure (secondary, tertiary and quaternary), local coiling and folding of the polypeptide

chain(s) give the molecule its final functional shape. All of these levels, however, derive from the protein's primary structure—that is, the precise location of specific amino acids in the polypeptide chain. The properties associated with a precise sequence of amino acids determine how the protein can twist and fold, thus adopting a specific stable structure that distinguishes it from every other protein.

Primary structure is established by covalent bonds. The next level of protein structure makes use of weaker hydrogen bonds.

The secondary structure of a protein requires hydrogen bonding

A protein's **secondary structure** consists of regular, repeated spatial patterns in different regions of a polypeptide chain. There are two basic types of secondary structure, both determined by hydrogen bonding between the amino acids that make up the primary structure, the α helix and the β pleated sheet.

THE α HELIX The α (**alpha**) **helix** is a right-handed coil that turns in the same direction as a standard wood screw (Figure 3.7B). The R groups extend outward from the peptide backbone of the helix. The coiling results from hydrogen bonds that form between the δ^+ hydrogen of the N—H of one amino acid and the δ^- oxygen of the C=O of another. When this pattern of hydrogen bonding is established repeatedly over a segment of the protein, it stabilizes the coil.

THE β PLEATED SHEET A β (**beta**) **pleated sheet** is formed from two or more polypeptide chains that are almost completely extended and aligned. The sheet is stabilized by hydrogen bonds between the N—H groups on one chain and the C=O groups on the other (Figure 3.7C). A β pleated sheet may form between separate polypeptide chains, as in spider silk, or be-

tween different regions of a single polypeptide chain that is bent back on itself. Many proteins contain regions of both α helix and β pleated sheet in the same polypeptide chain.

The tertiary structure of a protein is formed by bending and folding

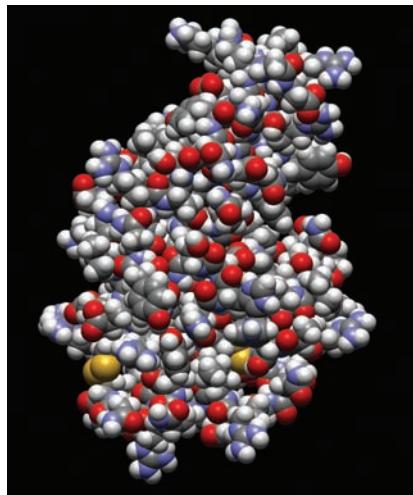
In many proteins, the polypeptide chain is bent at specific sites and then folded back and forth, resulting in the **tertiary structure** of the protein (Figure 3.7D). Although α helices and β pleated sheets contribute to the tertiary structure, usually only portions of the macromolecule have these secondary structures, and large regions consist of tertiary structure unique to a particular protein. Tertiary structure results in a macromolecule's definitive three-dimensional shape, often including a buried interior as well as a surface that is exposed to the environment.

The protein's exposed outer surfaces present functional groups capable of interacting with other molecules in the cell. These molecules might be other proteins (as happens in quaternary structure, as we will see below) or smaller chemical reactants (as in enzymes; see Section 7.4).

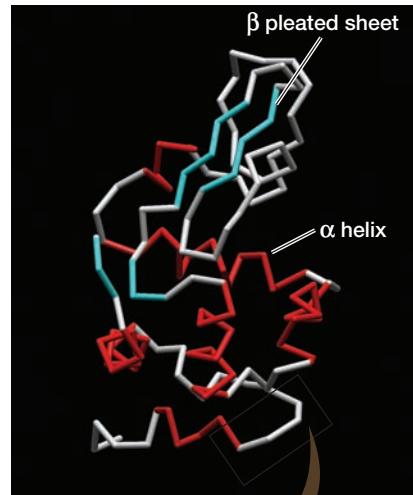
While hydrogen bonding between the N—H and C=O groups within and between chains is responsible for secondary structure, the interactions between R groups—the amino acid side chains—determine tertiary structure. We described the various strong and weak interactions between atoms in Section 2.2. Many of these interactions are involved in determining and maintaining tertiary structure.

3.8 Three Representations of Lysozyme Different molecular representations of a protein emphasize different aspects of its tertiary structure: surface features, sites of bends and folds, sites where alpha or beta structure predominate. These three representations of lysozyme are similarly oriented.

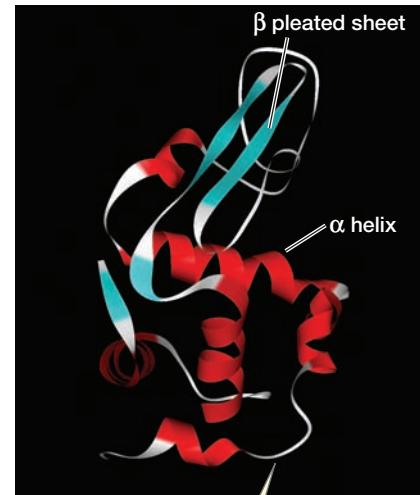
(A) Space-filling model



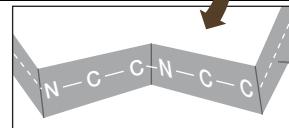
(B) Stick model



(C) Ribbon model



A realistic depiction of lysozyme shows dense packing of its atoms.



The "backbone" of lysozyme consists of repeating N—C—C units of amino acids.

- Covalent *disulfide bridges* can form between specific cysteine side chains (see Figure 3.5), holding a folded polypeptide in place.
- *Hydrogen bonds* between side chains also stabilize folds in proteins.
- *Hydrophobic side chains* can aggregate together in the interior of the protein, away from water, folding the polypeptide in the process.
- *van der Waals forces* can stabilize the close interactions between hydrophobic side chains.
- *Ionic bonds* can form between positively and negatively charged side chains, forming *salt bridges* between amino acids. Ionic bonds can also be buried deep within a protein, away from water.

A complete description of a protein's tertiary structure would specify the location of every atom in the molecule in three-dimensional space relative to all the other atoms. Such a description is available for the protein lysozyme (**Figure 3.8**).

The different ways of depicting the molecule have their uses. The space-filling model might be used to study how other molecules interact with specific sites and R groups on a protein's surface. The stick model emphasizes the sites where bends occur in order to make the folds of the polypeptide chain. The ribbon model, perhaps the most widely used, shows the different types of secondary structure and how they fold into the tertiary structure.

Remember that both secondary and tertiary structure derive from primary structure. If a protein is heated slowly, the heat energy will disrupt only the weak interactions, causing the secondary and tertiary structure to break down. The protein is then said to be **denatured**. But the protein can return to its normal tertiary structure when it cools, demonstrating that all the information needed to specify the unique shape of a protein is contained in its primary structure. This was first shown (using chemicals instead of heat to denature the protein) by biochemist Christian Anfinsen for the protein ribonuclease (**Figure 3.9**).

The quaternary structure of a protein consists of subunits

Many functional proteins contain two or more polypeptide chains, called *subunits*, each of them folded into its own unique tertiary structure. The protein's **quaternary structure** results from the ways in which these subunits bind together and interact (**Figure 3.7E**).

The models of hemoglobin in **Figure 3.10** illustrate quaternary structure. Hydrophobic interactions, van der Waals forces, hydrogen bonds, and ionic bonds all help hold the four subunits together to form a hemoglobin molecule. However, the weak nature of these forces permits small changes in the quaternary structure to aid the

INVESTIGATING LIFE

3.9 Primary Structure Specifies Tertiary Structure

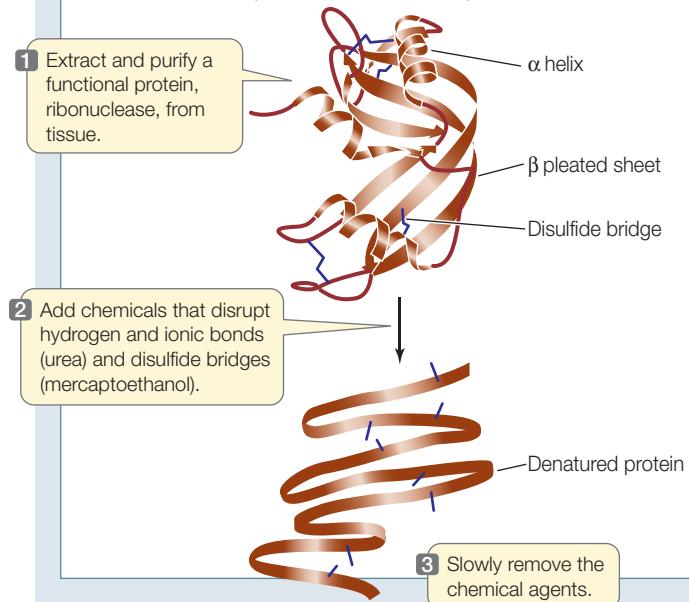
Using the protein ribonuclease, Christian Anfinsen showed that proteins spontaneously fold into a functionally correct three-dimensional configuration. As long as the primary structure is not disrupted, the information for correct folding under the right conditions is retained.

HYPOTHESIS

Under controlled conditions that simulate normal cellular environment in the laboratory, the primary structure of a denatured protein can reestablish the protein's three-dimensional structure.

METHOD

Chemically denature functional ribonuclease, disrupting disulfide bridges and other intramolecular interactions that maintain the protein's shape, so that only primary structure (i.e., the amino acid sequence) remains. Once denaturation is complete, remove the disruptive chemicals.



RESULTS

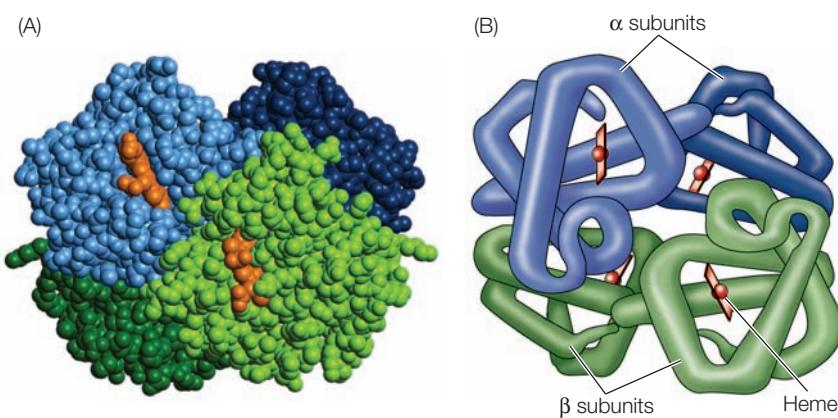
When the disruptive agents are removed, three-dimensional structure is restored and the protein once again is functional.



CONCLUSION

In normal cellular conditions, the primary structure of a protein specifies how it folds into a functional, three-dimensional structure.

3.10 Quaternary Structure of a Protein Hemoglobin consists of four folded polypeptide subunits that assemble themselves into the quaternary structure shown here. In these two graphic representations, each type of subunit is a different color. The heme groups contain iron and are the oxygen-carrying sites.



protein's function—which is to carry oxygen in red blood cells. As hemoglobin binds one O_2 molecule, the four subunits shift their relative positions slightly, changing the quaternary structure. Ionic bonds are broken, exposing buried side chains that enhance the binding of additional O_2 molecules. The quaternary structure changes again when hemoglobin releases its O_2 molecules to the cells of the body.

Shape and surface chemistry contribute to protein function

The shape and structure of a protein allow specific sites on its exposed surface to bind noncovalently to another molecule, which may be large or small. The binding is said to be specific because only certain compatible chemical groups will bind to one another. The specificity of protein binding depends on two general properties of the protein: its shape, and the chemistry of its exposed surface groups.

- **Shape.** When a small molecule collides with and binds to a much larger protein, it is like a baseball being caught by a catcher's mitt: the mitt has a shape that binds to the ball and fits around it. Just as a hockey puck or a ping-pong ball does not fit a baseball catcher's mitt, a given molecule will not bind to a protein unless there is a general "fit" between their two three-dimensional shapes.
- **Chemistry.** The exposed amino acid R groups on the surface of a protein permit chemical interactions with other substances (Figure 3.11). Three types of interactions may be involved: ionic, hydrophobic, and hydrogen bonding. Many important functions of proteins involve interactions between exposed-surface R groups and other molecules.

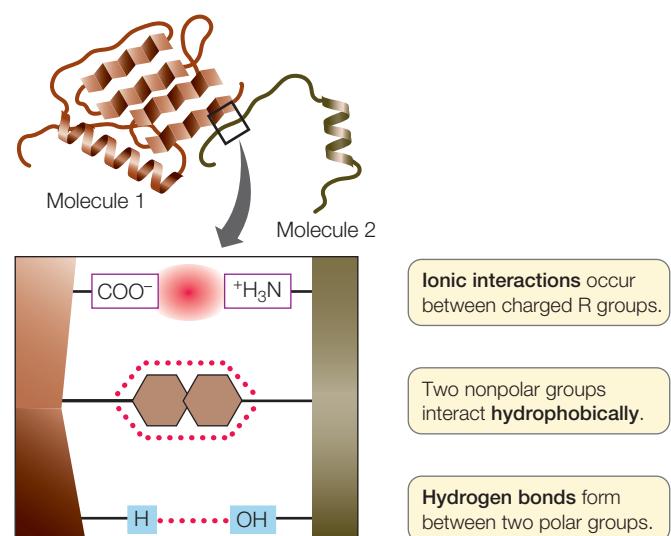
Environmental conditions affect protein structure

Because it is determined by weak forces, the three-dimensional structure of proteins is influenced by environmental conditions. Conditions that would not break covalent bonds can disrupt the

weaker, noncovalent interactions that determine secondary and tertiary structure. Such alterations may affect a protein's shape and thus its function. Various conditions can alter the weak, noncovalent interactions:

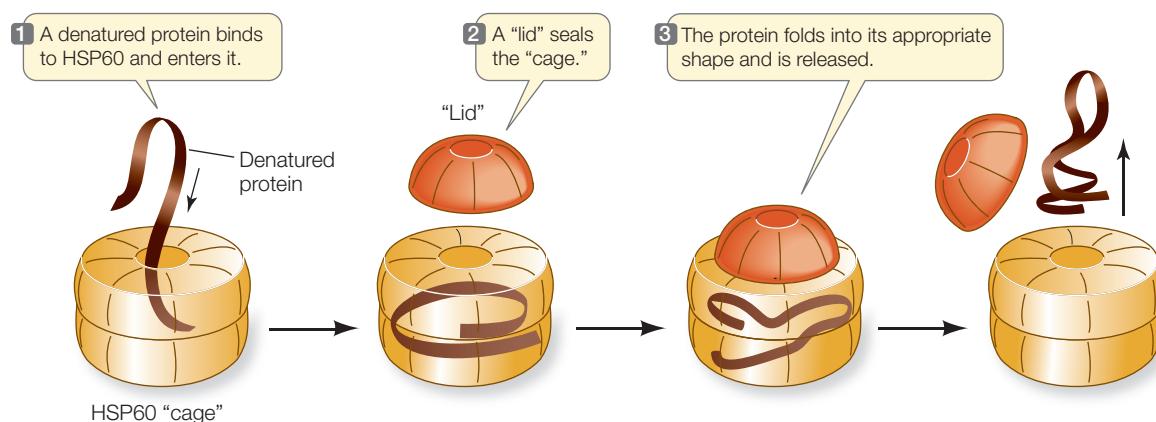
- *Increases in temperature* cause more rapid molecular movements and thus can break hydrogen bonds and hydrophobic interactions.
- *Alterations in pH* can change the pattern of ionization of exposed carboxyl and amino groups in the R groups of amino acids, thus disrupting the pattern of ionic attractions and repulsions.
- *High concentrations of polar substances* such as urea can disrupt the hydrogen bonding that is crucial to protein structure. This was used in the experiment on reversible protein denaturation shown in Figure 3.9.
- *Nonpolar substances* may also disrupt normal protein structure in cases where hydrophobic groups are essential to maintain the structure.

Denaturation can be irreversible when amino acids that were buried in the interior of the protein become exposed at the surface, and vice versa, causing a new structure to form or different molecules to bind to the protein. Boiling an egg denatures its proteins and is, as you know, not reversible.



3.11 Noncovalent Interactions Between Proteins and Other Molecules

Noncovalent interactions allow a protein (brown) to bind tightly to another molecule (green) with specific properties. Noncovalent interactions also allow regions within the same protein to interact with one another.



3.12 Chaperones Protect Proteins from Inappropriate Binding

Chaperone proteins surround new or denatured proteins and prevent them from binding to the wrong substance. Heat shock proteins such as HSP60, whose actions are illustrated here, are one class of chaperone proteins.

Molecular chaperones help shape proteins

Because of their specific shapes and the exposure of chemical groups on their surfaces, proteins can bind specific substances. Within a living cell, a polypeptide chain is sometimes in danger of binding the wrong substance. Two important examples of such a situation are:

- Following denaturation: Inappropriate environmental conditions in a cell, such as elevated temperature, can cause the denatured protein to re-fold incorrectly.
- Just after a protein is made: When a protein has not yet folded completely, it can present a surface that binds the wrong molecule.

In these cases, change may be irreversible. Eukaryotic cells have a special class of proteins that act to counteract threats to three-dimensional structure. Proteins in this class, called **chaperones**, act as molecular caretakers for other proteins. Like the chaperones at a high school dance, they prevent inappropriate interactions and enhance the appropriate ones.

Molecular chaperones were discovered by accident in 1962, when the temperature of an incubator holding fruit flies was accidentally turned up. Italian geneticist Ferruccio Ritossa noticed that this "heat shock" did not kill the flies. Instead, there was enhanced synthesis of a set of proteins that were later described as chaperones. They bound to many target proteins in the fruit fly cells and kept them from being denatured, and in some cases facilitated the correct refolding of proteins.

The general class of stress-induced chaperone proteins is called the **heat shock proteins (HSPs)**, after this discovery. HSPs are made by most eukaryotic cells, and many enhance protein folding in addition to their protective role during periods of stress. As an example, HSP60 forms a cage that sucks a protein in, causes it to fold into the correct shape, and then releases it (Figure 3.12). Tumors make abundant HSPs, possibly to stabilize proteins important in the cancer process, and so HSP-inhibiting drugs are being designed. In some clinical situations, treatment with these inhibitors results in the inappropriate folding of tumor-cell proteins, causing the tumors to stop growing and even disappear.

3.2 RECAP

Proteins are polymers of amino acids. The sequence of amino acids in a protein determines its primary structure. Secondary, tertiary, and quaternary structures arise through interactions between the amino acids. A protein's three-dimensional shape and exposed chemical groups establish binding specificity for other substances.

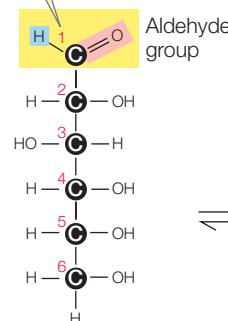
- What are the attributes of an amino acid's R group that would make it hydrophobic? Hydrophilic? See pp. 42–43 and Table 3.1
- Sketch and explain how two amino acids link together to form a peptide linkage. See p. 44 and Figure 3.6
- What are the four levels of protein structure and how are they all ultimately determined by the protein's primary structure (i.e., its amino acid sequence)? See pp. 44–48 and Figure 3.7
- How do environmental factors such as temperature and pH affect the weak interactions that give a protein its specific shape and function? See p. 48

The seemingly infinite number of protein configurations made possible by the biochemical properties of the 20 amino acids has driven the evolution of life's diversity. The linkage configurations of sugar monomers (monosaccharides) drives the structure of the next group of macromolecules, the carbohydrates that provide energy for life.

3.3 What Are the Chemical Structures and Functions of Carbohydrates?

Carbohydrates are a large group of molecules that all have a similar atomic composition but differ greatly in size, chemical properties, and biological functions. Carbohydrates have the general formula $C_n(H_2O)_n$, which makes them appear as hydrates of carbon (association between water molecules and carbon in the ratio $C_1H_2O_1$), hence their name. When their molecular structures are examined, the linked carbon atoms are seen to be bonded with hydrogen atoms ($-H$) and hydroxyl groups

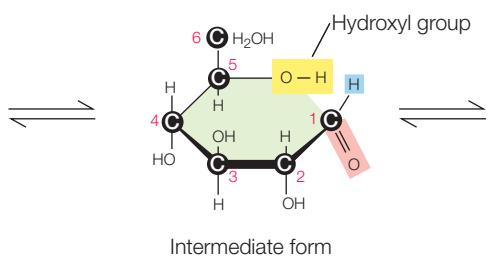
The numbers in red indicate the standard convention for numbering the carbons.



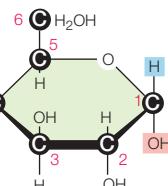
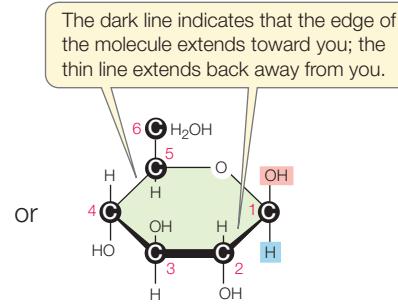
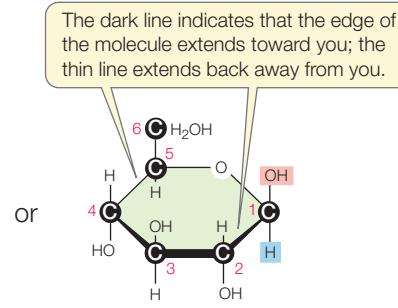
Straight-chain form

The straight-chain form of glucose has an aldehyde group at carbon 1.

3.13 From One Form of Glucose to the Other All glucose molecules have the formula $C_6H_{12}O_6$, but their structures vary. When dissolved in water, the α and β “ring” forms of glucose interconvert. The convention used here for numbering the carbon atoms is standard in biochemistry.



Intermediate form

 α -D-glucose β -D-glucose

or

Depending on the orientation of the aldehyde group when the ring closes, either of two molecules— α -D-glucose or β -D-glucose—forms.

yourBioPortal.com

GO TO Web Activity 3.3 • Forms of Glucose

($-\text{OH}$), the components of water. Carbohydrates have three major biochemical roles:

- They are a source of stored energy that can be released in a form usable by organisms.
- They are used to transport stored energy within complex organisms.
- They serve as *carbon skeletons* that can be rearranged to form new molecules.

Some carbohydrates are relatively small, with molecular weights of less than 100 Da. Others are true macromolecules, with molecular weights in the hundreds of thousands.

There are four categories of biologically important carbohydrates:

- **Monosaccharides** (*mono*, “one”; *saccharide*, “sugar”), such as glucose, ribose, and fructose, are *simple sugars*. They are the monomers from which the larger carbohydrates are constructed.
- **Disaccharides** (*di*, “two”) consist of two monosaccharides linked together by covalent bonds. The most familiar is sucrose, which is made up of covalently bonded glucose and fructose molecules.
- **Oligosaccharides** (*oligo*, “several”) are made up of several (3–20) monosaccharides.
- **Polysaccharides** (*poly*, “many”), such as starch, glycogen, and cellulose, are polymers made up of hundreds or thousands of monosaccharides.

Monosaccharides are simple sugars

All living cells contain the monosaccharide **glucose**; it is the familiar “blood sugar,” used to transport energy in humans. Cells use glucose as an energy source, breaking it down through a series of reactions that release stored energy and produce water and carbon dioxide; this is a cellular form of the combustion reaction described in Chapter 2.

Glucose exists in straight chains and in ring forms. The ring forms predominate in virtually all biological circumstances because they are more stable under physiological conditions. There are two versions of glucose ring, called α - and β -glucose, which differ only in the orientation of the $-\text{H}$ and $-\text{OH}$ attached to carbon 1 (Figure 3.13). The α and β forms interconvert and exist in equilibrium when dissolved in water.

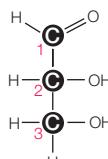
Different monosaccharides contain different numbers of carbons. Some monosaccharides are structural isomers, with the same kinds and numbers of atoms, but in different arrangements (Figure 3.14). Such seemingly small structural changes can significantly alter properties. Most of the monosaccharides in living systems belong to the D (right-handed) series of isomers.

Pentoses (*pente*, “five”) are five-carbon sugars. Two pentoses are of particular biological importance: the backbones of the nucleic acids RNA and DNA contain ribose and deoxyribose, respectively (see Section 4.1). These two pentoses are not isomers of each other; rather, one oxygen atom is missing from carbon 2 in deoxyribose (*de-*, “absent”). The absence of this oxygen atom is an important distinction between RNA and DNA.

The **hexoses** (*hex*, “six”), a group of structural isomers, all have the formula $C_6H_{12}O_6$. Included among the hexoses are glucose, fructose (so named because it was first found in fruits), mannose, and galactose.

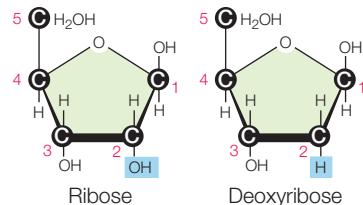
Glycosidic linkages bond monosaccharides

The disaccharides, oligosaccharides, and polysaccharides are all constructed from monosaccharides that are covalently bonded together by condensation reactions that form **glycosidic linkages**. A single glycosidic linkage between two monosaccharides forms a disaccharide. For example, sucrose—common table sugar in the human diet and a major disaccharide in plants—is a disaccharide formed from a glucose and a fructose molecule.

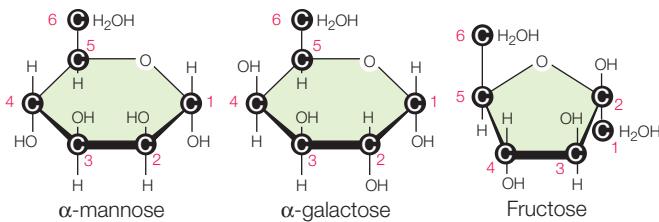
Three-carbon sugar

Glyceraldehyde is the smallest monosaccharide and exists only as the straight-chain form.

Glyceraldehyde

Five-carbon sugars (pentoses)

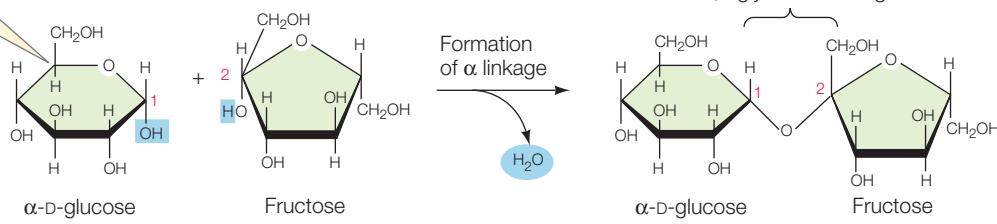
Ribose and deoxyribose each have five carbons, but very different chemical properties and biological roles.

Six-carbon sugars (hexoses)

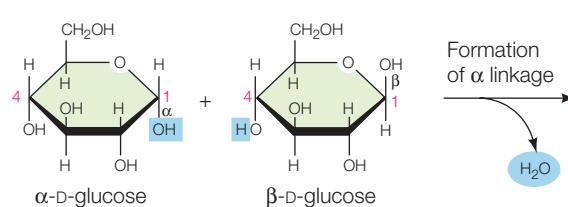
These hexoses are structural isomers. All have the formula $C_6H_{12}O_6$, but each has distinct biochemical properties.

The presence of a carbon atom (C) at a junction such as this is implied.

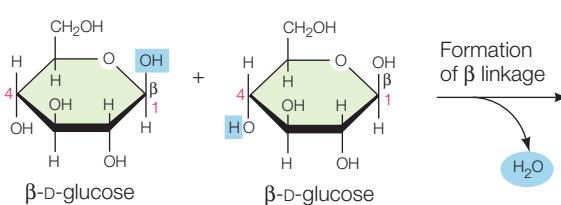
In sucrose, glucose and fructose are linked by an α -1,2 glycosidic linkage.



Maltose is produced when an α -1,4 glycosidic linkage forms between two glucose molecules. The hydroxyl group on carbon 1 of one D-glucose in the α (down) position reacts with the hydroxyl group on carbon 4 of the other glucose.



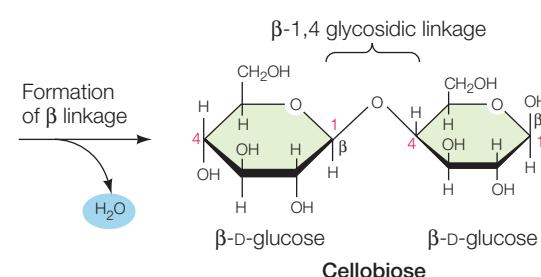
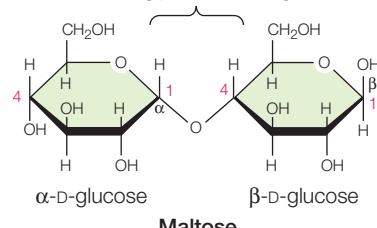
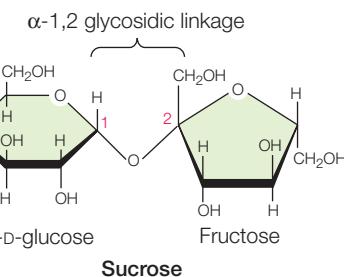
In cellobiose, two glucoses are linked by a β -1,4 glycosidic linkage.



The disaccharides maltose and cellobiose are made from two glucose molecules (Figure 3.15). Maltose and cellobiose are structural isomers, both having the formula $C_{12}H_{22}O_{11}$. However, they have different chemical properties and are recognized by different enzymes in biological tissues. For example, maltose can be hydrolyzed into its monosaccharides in the human body, whereas cellobiose cannot.

Oligosaccharides contain several monosaccharides bound by glycosidic linkages at various sites. Many oligosaccharides have additional functional groups, which give them special properties. Oligosaccharides are often covalently bonded to proteins and lipids on the outer cell surface, where they serve as recognition signals. The different human blood groups (for example, the ABO blood types) get their specificity from oligosaccharide chains.

3.15 Disaccharides Form by Glycosidic Linkages Glycosidic linkages between two monosaccharides can create many different disaccharides. Which disaccharide is formed depends on which monosaccharides are linked; on the site of linkage (i.e., which carbon atoms are involved); and on the form (α or β) of the linkage.



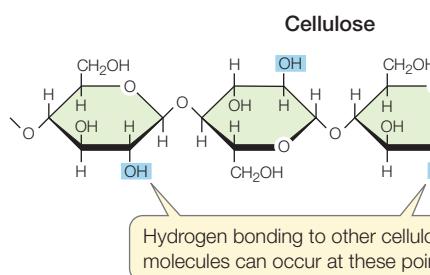
Polysaccharides store energy and provide structural materials

Polysaccharides are large (sometimes gigantic) polymers of monosaccharides connected by glycosidic linkages (Figure 3.16). In contrast to proteins, polysaccharides are not necessarily linear chains of monomers. Each monomer unit has several sites that may be capable of forming glycosidic linkages, and thus branched molecules are possible.

STARCH Starches comprise a family of giant molecules of broadly similar structure. While all starches are polysaccharides of glucose with α -glycosidic linkages (α -1,4 and α -1,6 glycosidic

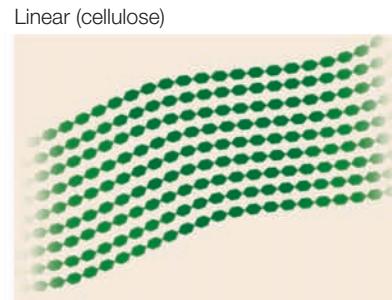
bonds; Figure 3.16A), the different starches can be distinguished by the amount of branching that occurs at carbons 1 and 6 (Figure 3.16B). Starch is the principal energy storage compound of plants. Some plant starches, such as amylose, are unbranched; others are moderately branched (amylopectin, for example). Starch readily binds water. When that water is removed, however, hydrogen bonds tend to form between the unbranched polysaccharide chains, which then aggregate, as in the large starch grains observed in the storage material of plant seeds (see Figure 3.16C).

(A) Molecular structure

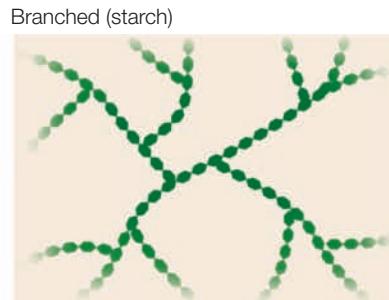


Cellulose is an unbranched polymer of glucose with β -1,4 glycosidic linkages that are chemically very stable.

(B) Macromolecular structure

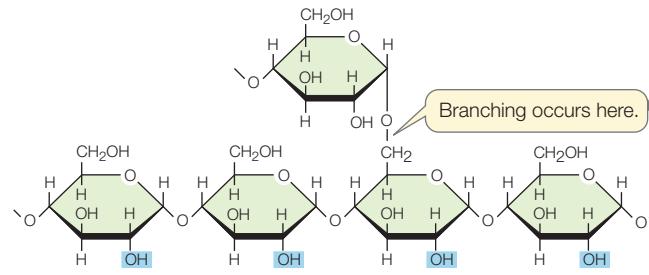


Parallel cellulose molecules form hydrogen bonds, resulting in thin fibrils.



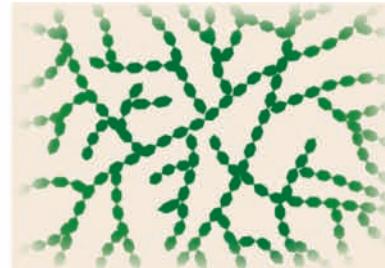
Branching limits the number of hydrogen bonds that can form in starch molecules, making starch less compact than cellulose.

Starch and glycogen



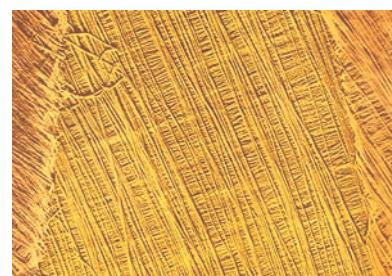
Glycogen and starch are polymers of glucose with α -1,4 glycosidic linkages. α -1,6 glycosidic linkages produce branching at carbon 6.

Highly branched (glycogen)

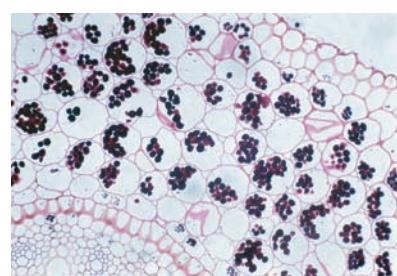


The high amount of branching in glycogen makes its solid deposits more compact than starch.

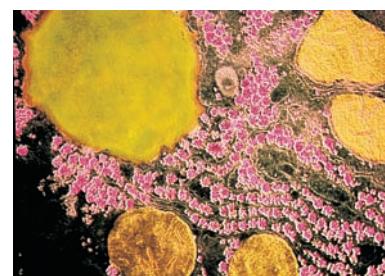
(C) Polysaccharides in cells



Layers of cellulose fibrils, as seen in this scanning electron micrograph, give plant cell walls great strength.



Within these plant cells, starch deposits (dyed purple in this micrograph) have a granular shape.



The pink-stained granules in this electron micrograph are glycogen deposits in the human liver.

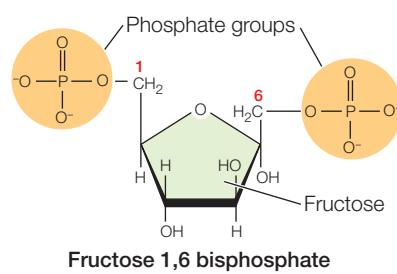
GLYCOGEN Glycogen is a water-insoluble, highly branched polymer of glucose. It stores glucose in liver and muscle, serving as an energy storage compound for animals as starch does for plants. Both glycogen and starch are readily hydrolyzed into glucose monomers, which in turn can be broken down to liberate their stored energy.

But if it is glucose that is needed for fuel, why store it in the form of glycogen? The reason is that 1,000 glucose molecules would exert 1,000 times the *osmotic pressure* of a single glycogen molecule, causing water to enter the cells (see Section 6.3). If it were not for polysaccharides, many organisms would expend a lot of energy expelling excess water from their cells.

CELLULOSE As the predominant component of plant cell walls, cellulose is by far the most abundant organic compound on Earth. Like starch and glycogen, cellulose is a polysaccharide of glucose, but its individual monosaccharides are connected by β - rather than by α -glycosidic linkages. Starch is easily degraded by the actions of chemicals or enzymes. Cellulose, however, is chemically more stable because of its β -glycosidic linkages. Thus, whereas starch is easily broken down to supply glucose for energy-producing reactions, cellulose is an excellent structural material that can withstand harsh environmental conditions without substantial change.

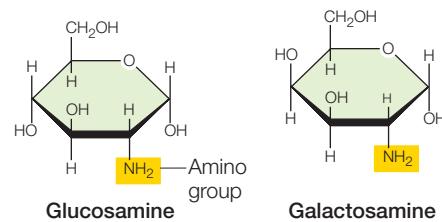
(A) Sugar phosphate

Fructose 1,6 bisphosphate is involved in the reactions that liberate energy from glucose. (The numbers in its name refer to the carbon sites of phosphate bonding; *bis-* indicates that two phosphates are present.)



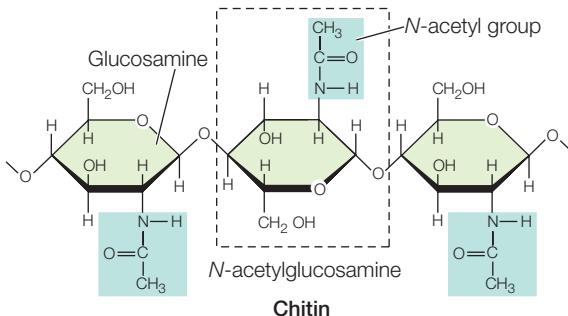
(B) Amino sugars

The monosaccharides glucosamine and galactosamine are amino sugars with an amino group in place of a hydroxyl group.



(C) Chitin

Chitin is a polymer of *N*-acetylglucosamine; *N*-acetyl groups provide additional sites for hydrogen bonding between the polymers.



Chemically modified carbohydrates contain additional functional groups

Some carbohydrates are chemically modified by the addition of functional groups, such as phosphate and amino groups (Figure 3.17). For example, carbon 6 in glucose may be oxidized from $-\text{CH}_2\text{OH}$ to a carboxyl group ($-\text{COOH}$), producing glucuronic acid. Or a phosphate group may be added to one or more of the $-\text{OH}$ sites. Some of the resulting sugar phosphates, such as fructose 1,6-bisphosphate, are important intermediates in cellular energy reactions, which will be discussed in Chapter 9.

When an amino group is substituted for an $-\text{OH}$ group, *amino sugars*, such as glucosamine and galactosamine, are produced. These compounds are important in the extracellular matrix (see Section 5.4), where they form parts of glycoproteins, which are molecules involved in keeping tissues together. Galactosamine is a major component of cartilage, the material that forms caps on the ends of bones and stiffens the ears and nose. A derivative of glucosamine is present in the polymer *chitin*, the principal structural polysaccharide in the external skeletons of insects and many crustaceans (e.g., crabs and lobsters) and a component of the cell walls of fungi. Because these organisms are among the most abundant eukaryotes on Earth, chitin rivals cellulose as one of the most abundant substances in the living world.

3.17 Chemically Modified Carbohydrates Added functional groups can modify the form and properties of a carbohydrate.

Galactosamine is an important component of cartilage, a connective tissue in vertebrates.



The external skeletons of insects are made up of chitin.



3.3 RECAP

Carbohydrates are composed of carbon, hydrogen, and oxygen in the general ratio of 1:2:1. They provide energy and structure to cells and are precursors of numerous important biological molecules. Monosaccharide monomers can be connected by glycosidic linkages to form disaccharides, oligosaccharides, and polysaccharides.

- Draw the chemical structure of a disaccharide formed by two monosaccharides. **See Figure 3.15**
- What qualities of the polysaccharides starch and glycogen make them useful for energy storage? **See pp. 52–53 and Figure 3.16**
- From looking at the cellulose molecules in Figure 3.16A, can you see where a large number of hydrogen bonds are present in the linear structure of cellulose shown in Figure 3.16B? Why is this structure so strong?

We have seen how amino acid monomers form protein polymers and how sugar monomers form the polymers of carbohydrates. Now we will look at the lipids, which are unique among the four classes of large biological molecules in that they are not, strictly speaking, polymers.

3.4 What Are the Chemical Structures and Functions of Lipids?

Lipids—colloquially called *fats*—are hydrocarbons that are insoluble in water because of their many nonpolar covalent bonds. As we saw in Section 2.2, nonpolar hydrocarbon molecules are hydrophobic and preferentially aggregate among themselves, away from water (which is polar). When nonpolar hydrocarbons are sufficiently close together, weak but additive van der Waals forces hold them together. The huge macromolecular aggregations that can form are not polymers in a strict chemical sense, because the individual lipid molecules are not covalently bonded. With this understanding, it is still useful to consider aggregations of individual lipids as a different sort of polymer.

There are several different types of lipids, and they play a number of roles in living organisms:

- Fats and oils store energy.
- Phospholipids play important structural roles in cell membranes.
- Carotenoids and chlorophylls help plants capture light energy.
- Steroids and modified fatty acids play regulatory roles as hormones and vitamins.

- Fat in animal bodies serves as thermal insulation.
- A lipid coating around nerves provides electrical insulation.
- Oil or wax on the surfaces of skin, fur, and feathers repels water.

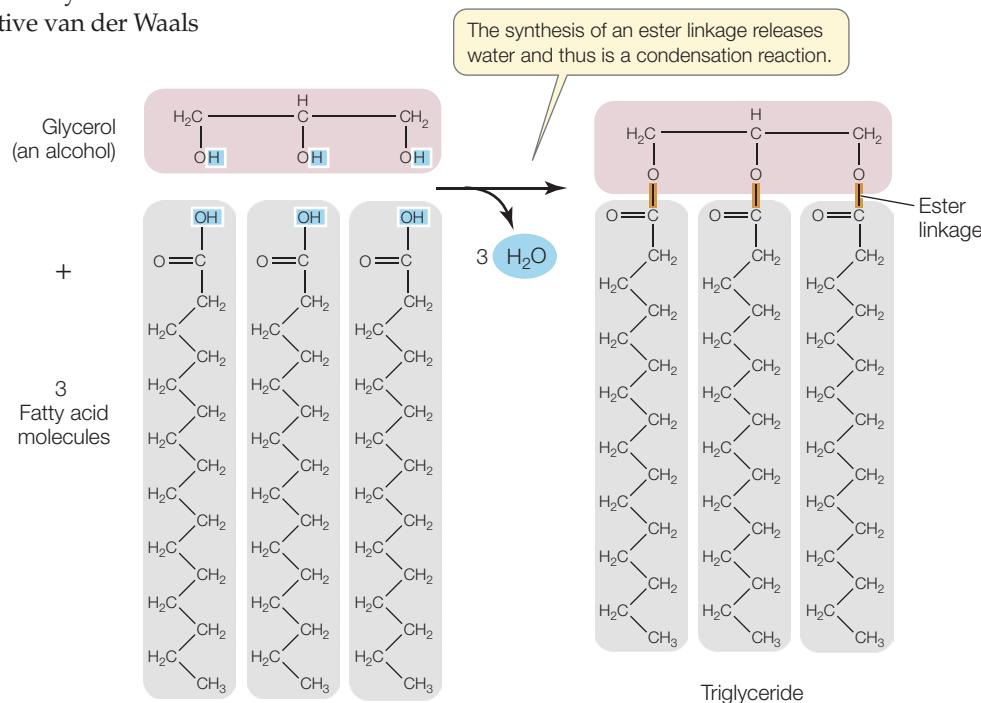
Fats and oils are hydrophobic

Chemically, fats and oils are *triglycerides*, also known as *simple lipids*. Triglycerides that are solid at room temperature (around 20°C) are called **fats**; those that are liquid at room temperature are called **oils**. Triglycerides are composed of two types of building blocks: *fatty acids* and *glycerol*. **Glycerol** is a small molecule with three hydroxyl ($-\text{OH}$) groups (thus it is an alcohol). A **fatty acid** is made up of a long nonpolar hydrocarbon chain and a polar carboxyl group ($-\text{COOH}$). These chains are very hydrophobic, with their abundant C—H and C—C bonds, which have low electronegativity and are nonpolar (see Section 2.2).

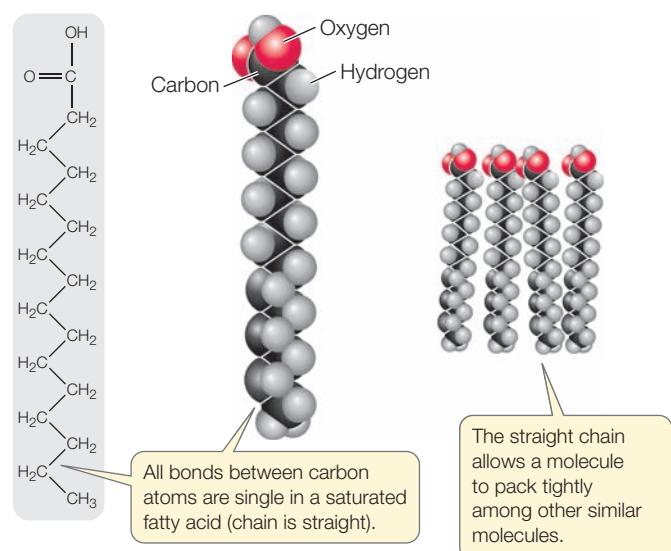
A **triglyceride** contains three fatty acid molecules and one molecule of glycerol. Synthesis of a triglyceride involves three condensation (dehydration) reactions. In each reaction, the carboxyl group of a fatty acid bonds with a hydroxyl group of glycerol, resulting in a covalent bond called an **ester linkage** and the release of a water molecule (**Figure 3.18**). The three fatty acids in a triglyceride molecule need not all have the same hydrocarbon chain length or structure; some may be saturated fatty acids, while others may be unsaturated:

- In **saturated fatty acids**, all the bonds between the carbon atoms in the hydrocarbon chain are single bonds—there are no double bonds. That is, all the bonds are saturated with

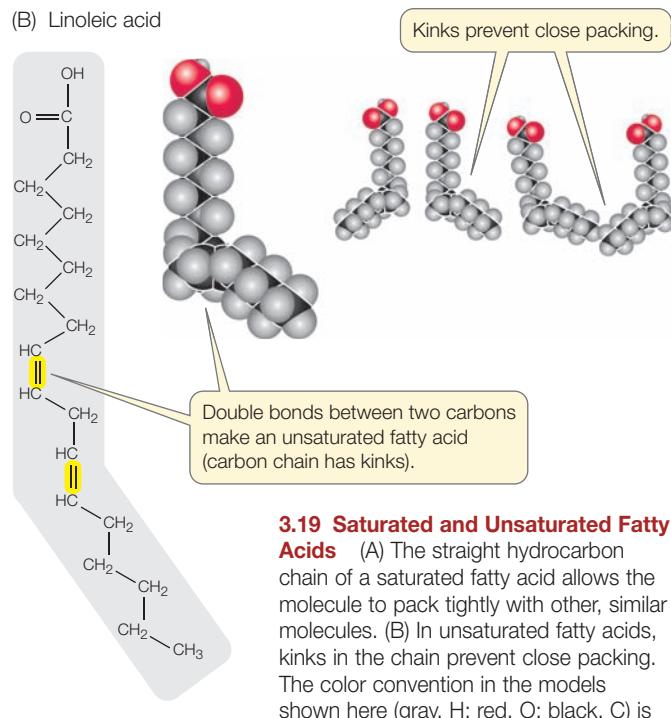
3.18 Synthesis of a Triglyceride In living things, the reaction that forms a triglyceride is more complex, but the end result is the same as shown here.



(A) Palmitic acid



(B) Linoleic acid



3.19 Saturated and Unsaturated Fatty Acids (A) The straight hydrocarbon chain of a saturated fatty acid allows the molecule to pack tightly with other, similar molecules. (B) In unsaturated fatty acids, kinks in the chain prevent close packing. The color convention in the models shown here (gray, H; red, O; black, C) is commonly used.

hydrogen atoms (**Figure 3.19A**). These fatty acid molecules are relatively rigid and straight, and they pack together tightly, like pencils in a box.

- In **unsaturated fatty acids**, the hydrocarbon chain contains one or more double bonds. Linoleic acid is an example of a *polyunsaturated* fatty acid that has two double bonds near the middle of the hydrocarbon chain, which causes kinks in the molecule (**Figure 3.19B**). Such kinks prevent the unsaturated fat molecules from packing together tightly.

The kinks in fatty acid molecules are important in determining the fluidity and melting point of a lipid. The triglycerides

of animal fats tend to have many long-chain saturated fatty acids, packed tightly together; these fats are usually solids at room temperature and have a high melting point. The triglycerides of plants, such as corn oil, tend to have short or unsaturated fatty acids. Because of their kinks, these fatty acids pack together poorly and have a low melting point, and these triglycerides are usually liquids at room temperature.

Fats are excellent storehouses for chemical energy. As you will see in Chapter 9, when the C—H bond is broken, it releases significant energy that an organism can use for its own purposes, such as movement or building up complex molecules. On a per weight basis, broken-down fats yield more than twice as much energy as do degraded carbohydrates.

Phospholipids form biological membranes

We have mentioned the hydrophobic nature of the many C—C and C—H bonds in fatty acids. But what about the carboxyl functional group at the end of the molecule? When it ionizes and forms COO^- , it is strongly hydrophilic. So a fatty acid is a molecule with a hydrophilic end and a long hydrophobic tail. It has two opposing chemical properties; the technical term for this is **amphipathic**. This explains what happens when oil (fatty acid) and water mix: the fatty acids orient themselves so that their polar ends face outward (i.e., toward the water) and their nonpolar tails face inward (away from water). Although no covalent bonds link individual lipids in large aggregations, such stable aggregations form readily in aqueous conditions. So these large lipid structures can be considered a different kind of macromolecule.

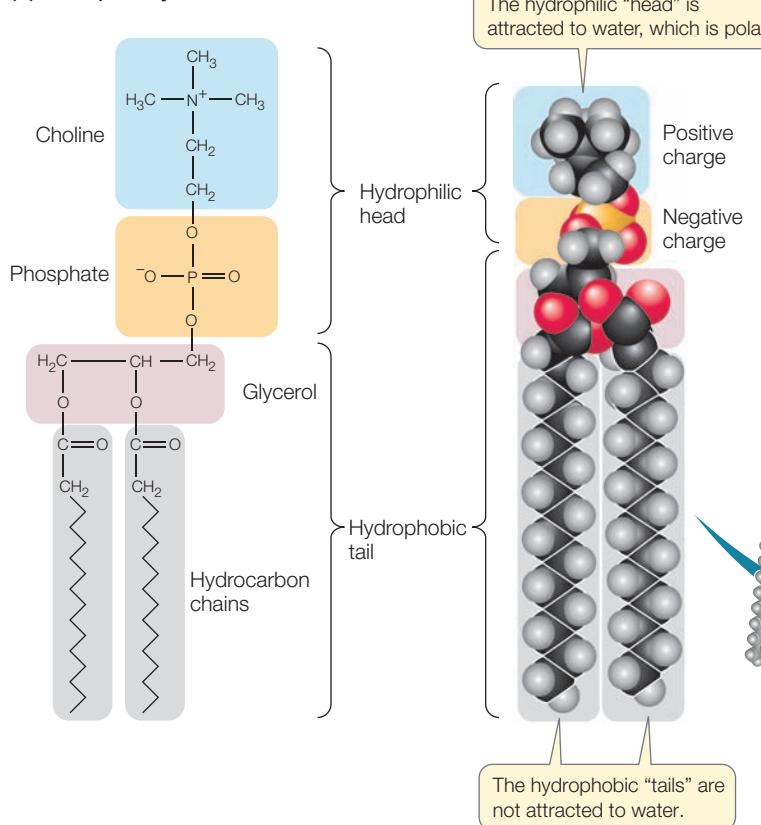
Like triglycerides, **phospholipids** contain fatty acids bound to glycerol by ester linkages. In phospholipids, however, any one of several phosphate-containing compounds replaces one of the fatty acids, giving these molecules amphipathic properties—that is properties of both water soluble and water insoluble molecules (**Figure 3.20A**). The phosphate functional group has a negative electric charge, so this portion of the molecule is hydrophilic, attracting polar water molecules. But the two fatty acids are hydrophobic, so they tend to avoid water and aggregate together or with other hydrophobic substances.

In an aqueous environment, phospholipids line up in such a way that the nonpolar, hydrophobic “tails” pack tightly together and the phosphate-containing “heads” face outward, where they interact with water. The phospholipids thus form a **bilayer**: a sheet two molecules thick, with water excluded from the core (**Figure 3.20B**). Biological membranes have this kind of **phospholipid bilayer** structure, and we will devote Chapter 6 to their biological functions.

Lipids have roles in energy conversion, regulation, and protection

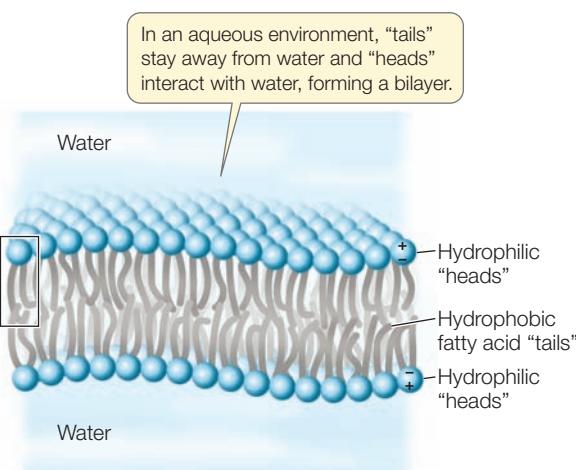
In the previous section, we focused on lipids involved in energy storage and cell structure, whose molecular structures are variations on the glycerol-fatty acid structure. However, there are other nonpolar and amphipathic lipids that are not based on this structure.

(A) Phosphatidylcholine

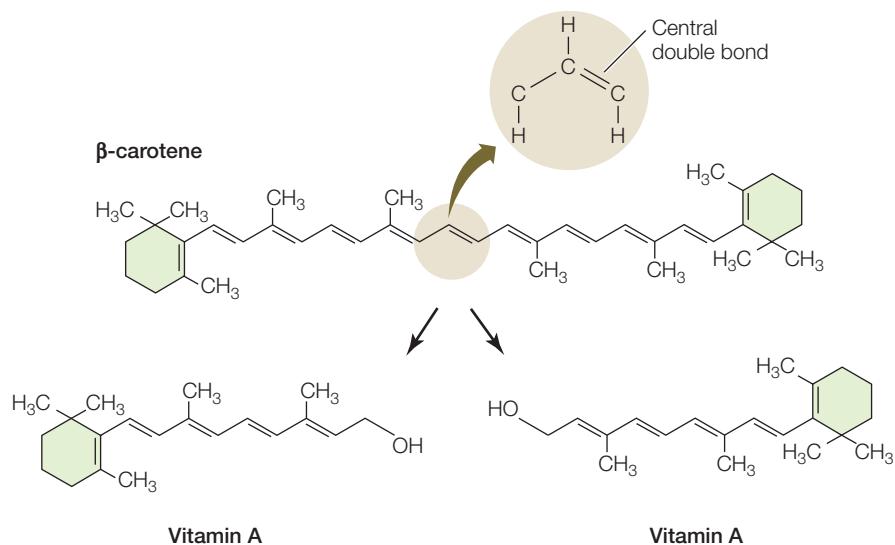


3.20 Phospholipids (A) Phosphatidylcholine (lecithin) demonstrates the structure of a phospholipid molecule. In other phospholipids, the amino acid serine, the sugar alcohol inositol, or other compounds replace choline. (B) In an aqueous environment, hydrophobic interactions bring the “tails” of phospholipids together in the interior of a bilayer. The hydrophilic “heads” face outward on both sides of the bilayer, where they interact with the surrounding water molecules.

(B) Phospholipid bilayer



CAROTENOIDS The carotenoids are a family of light-absorbing pigments found in plants and animals. Beta-carotene (β -carotene) is one of the pigments that traps light energy in leaves during photosynthesis. In humans, a molecule of β -carotene can be broken down into two vitamin A molecules (Figure 3.21), from which we make the pigment *cis*-retinal, which is required for vision. Carotenoids are responsible for the colors of carrots, tomatoes, pumpkins, egg yolks, and butter.

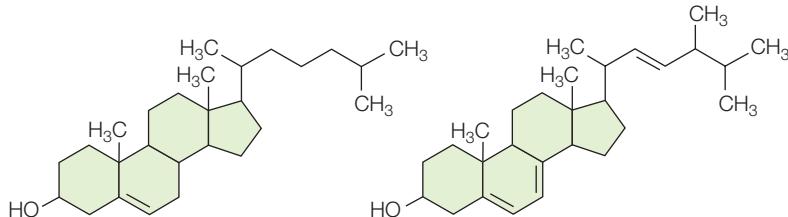


STEROIDS The steroids are a family of organic compounds whose multiple rings share carbons (Figure 3.22). The steroid cholesterol is an important constituent of membranes. Other steroids function as hormones, chemical signals that carry messages from one part of the body to another (see Chapter 41). Cholesterol is synthesized in the liver and is the starting material for making testosterone and other steroid hormones, such as estrogen.

VITAMINS Vitamins are small molecules that are not synthesized by the human body and so must be acquired from the diet (see

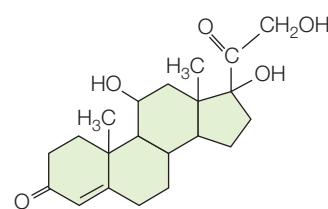
Chapter 50). For example, vitamin A is formed from the β -carotene found in green and yellow vegetables (see Figure 3.21). In humans, a deficiency of vitamin A leads to dry skin, eyes, and internal body surfaces, retarded growth and development, and night blindness, which is a diagnostic symptom for the deficiency. Vitamins D, E, and K are also lipids.

3.21 β -Carotene is the Source of Vitamin A The carotenoid β -carotene is symmetrical around its central double bond. When that bond is broken, two molecules of vitamin A are formed. The structural formula presented here is standard chemical shorthand for large organic molecules with many carbon atoms; it is simplified by omitting the C (indicating a carbon atom) at the intersections representing covalent bonds. The presence of hydrogen atoms (H) to fill all the available bonding sites on each C is assumed.

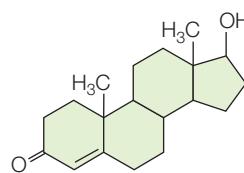


Cholesterol is a constituent of membranes and is the source of steroid hormones.

Vitamin D₂ can be produced in the skin by the action of light on a cholesterol derivative.

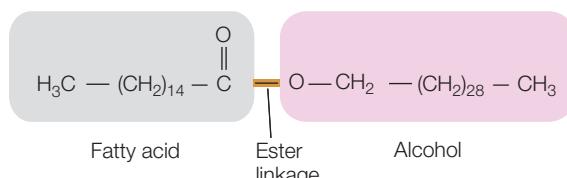


Cortisol is a hormone secreted by the adrenal glands.



Testosterone is a male sex hormone.

WAXES The sheen on human hair is more than cosmetic. Glands in the skin secrete a waxy coating that repels water and keeps the hair pliable. Birds that live near water have a similar waxy coating on their feathers. The shiny leaves of plants such as holly, familiar during winter holidays, also have a waxy coating. Finally, bees make their honeycombs out of wax. All waxes have the same basic structure: they are formed by an ester linkage between a saturated, long-chain fatty acid and a saturated, long-chain alcohol. The result is a very long molecule, with 40–60 CH₂ groups. For example, here is the structure of beeswax:



This highly nonpolar structure accounts for the impermeability of wax to water.

3.22 All Steroids Have the Same Ring Structure The steroids shown here, all important in vertebrates, are composed of carbon and hydrogen and are highly hydrophobic. However, small chemical variations, such as the presence or absence of a hydroxyl group, can produce enormous functional differences among these molecules.

3.4 RECAP

Lipids include both nonpolar and amphipathic molecules that are largely composed of carbon and hydrogen. They are important in energy storage, light absorption, regulation and biological structures. Cell membranes contain phospholipids, which are composed of hydrophobic fatty acids linked to glycerol and a hydrophilic phosphate group.

- Draw the molecular structures of fatty acids and glycerol and show how they are linked to form a triglyceride. **See p. 54 and Figure 3.18**
- What is the difference between fats and oils? **See p. 54**
- How does the polar nature of phospholipids result in their forming a bilayer? **See p. 55 and Figure 3.20**
- Why are steroids and some vitamins classified as lipids? **See p. 56**

All the types of molecules we have discussed in this chapter are found only in living organisms, but a final class of biological macromolecules has special importance to the living world. The function of the nucleic acids is nothing less than the transmission of life's "blueprint" to each new organism. This chapter showed the wonderful biochemical unity of life, a unity that implies all life has a common origin. Essential to this origin were the monomeric nucleotides and their polymers, nucleic acids. In the next chapter, we turn to the related topics of nucleic acids and the origin of life.

CHAPTER SUMMARY

3.1 What Kinds of Molecules Characterize Living Things?

SEE ANIMATED TUTORIAL 3.1

- Macromolecules** are **polymers** constructed by the formation of covalent bonds between smaller molecules called **monomers**. Macromolecules in living organisms include polysaccharides, proteins, and nucleic acids. Large lipid structures may also be considered macromolecules.
- Functional groups** are small groups of atoms that are consistently found together in a variety of different macromolecules. Functional groups have particular chemical properties that they

confer on any larger molecule of which they are a part. **Review Figure 3.1, WEB ACTIVITY 3.1**

- Structural and optical **isomers** have the same kinds and numbers of atoms, but differ in their structures and properties. **Review Figure 3.2**
- The many functions of macromolecules are directly related to their three-dimensional shapes, which in turn result from the sequences and chemical properties of their monomers.
- Monomers are joined by **condensation reactions**, which release a molecule of water for each bond formed. **Hydrolysis reactions** use water to break polymers into monomers. **Review Figure 3.4**

3.2 What Are the Chemical Structures and Functions of Proteins?

- The functions of proteins include support, protection, catalysis, transport, defense, regulation, and movement.
- Amino acids** are the monomers from which proteins are constructed. Four groups are attached to a central carbon atom: a hydrogen atom, an amino group, a carboxyl group, and a variable R group. The particular properties of each amino acids depend on its **side chain**, or **R group**, which may be charged, polar, or hydrophobic. **Review Table 3.1, WEB ACTIVITY 3.2**
- Peptide linkages**, also called peptide bonds, covalently link amino acids into polypeptide chains. These bonds form by condensation reactions between the carboxyl and amino groups. **Review Figure 3.6**
- The **primary structure** of a protein is the sequence of amino acids in the chain. This chain is folded into a **secondary structure**, which in different parts of the protein may form an **α helix** or a **β pleated sheet**. **Review Figure 3.7A–C**
- Disulfide bridges** and noncovalent interactions between amino acids cause polypeptide chains to fold into three-dimensional **tertiary structures** and allow multiple chains to interact in a **quaternary structure**. **Review Figure 3.7D,E**
- Heat, alterations in pH, or certain chemicals can all result in protein **denaturation**, which involves the loss of tertiary and/or secondary structure as well as biological function. **Review Figure 3.9**
- The specific shape and structure of a protein allows it to bind noncovalently to other molecules. **Review Figure 3.11**
- Chaperone proteins** enhance correct protein folding and prevent binding to inappropriate ligands. **Review Figure 3.12**

3.3 What Are the Chemical Structures and Functions of Carbohydrates?

- Carbohydrates** contain carbon bonded to hydrogen and oxygen atoms in a ratio of 1:2:1, or $(\text{CH}_2\text{O})_n$.

- Monosaccharides** are the monomers that make up carbohydrates. **Hexoses** such as **glucose** are six-carbon monosaccharides; **pentoses** have five carbons. **Review Figure 3.14, WEB ACTIVITY 3.3**
- Glycosidic linkages**, which have either an α or a β orientation in space, covalently link monosaccharides into larger units such as **disaccharides**, **oligosaccharides**, and **polysaccharides**. **Review Figure 3.15**
- Starch** stores energy in plants. Starch and **glycogen** are formed by α -glycosidic linkages between glucose monomers and are distinguished by the amount of branching they exhibit. They can be easily broken down to release stored energy. **Review Figure 3.16**
- Cellulose** is a very stable glucose polymer and is the principal structural component of plant cell walls.

3.4 What Are the Chemical Structures and Functions of Lipids?

- Fats and oils are **triglycerides**, composed of three fatty acids covalently bonded to a molecule of glycerol by ester linkages. **Review Figure 3.18**
- Saturated** fatty acids have a hydrocarbon chain with no double bonds. The hydrocarbon chains of **unsaturated** fatty acids have one or more double bonds that bend the chain, making close packing less possible. **Review Figure 3.19**
- Phospholipids** have a hydrophobic hydrocarbon “tail” and a hydrophilic phosphate “head”; that is, they are **amphipathic**. In water, the interactions of the tails and heads of phospholipids generate a **phospholipid bilayer**. The heads are directed outward, where they interact with the surrounding water. The tails are packed together in the interior of the bilayer, away from water. **Review Figure 3.20**
- Other lipids include vitamins A and D, steroids and plant pigments such as carotenoids.

SELF-QUIZ

- The most abundant molecule in the cell is
 - a carbohydrate.
 - a lipid.
 - a nucleic acid.
 - a protein.
 - water.
- All lipids are
 - triglycerides.
 - polar.
 - hydrophilic.
 - polymers of fatty acids.
 - more soluble in nonpolar solvents than in water.
- All carbohydrates
 - are polymers.
 - are simple sugars.
 - consist of one or more simple sugars.
 - are found in biological membranes.
 - are more soluble in nonpolar solvents than in water.
- Which of the following is *not* a carbohydrate?
 - Glucose
 - Starch
 - Cellulose
 - Hemoglobin
 - Deoxyribose
- All proteins
 - are enzymes.
 - consist of one or more polypeptide chains.
 - are amino acids.
 - have quaternary structures.
 - are more soluble in nonpolar solvents than in water.
- Which of the following statements about the primary structure of a protein is *not* true?
 - It may be branched.
 - It is held together by covalent bonds.
 - It is unique to that protein.
 - It determines the tertiary structure of the protein.
 - It is the sequence of amino acids in the protein.

7. The amino acid leucine
 - a. is found in all proteins.
 - b. cannot form peptide linkages.
 - c. has a hydrophobic side chain.
 - d. has a hydrophilic side chain.
 - e. is identical to the amino acid lysine.
8. The quaternary structure of a protein
 - a. consists of four subunits—hence the name quaternary.
 - b. is unrelated to the function of the protein.
 - c. may be either alpha or beta.
 - d. depends on covalent bonding among the subunits.
 - e. depends on the primary structures of the subunits.
9. The amphipathic nature of phospholipids is
 - a. determined by the fatty acid composition.
 - b. important in membrane structure.
 - c. polar but not nonpolar.
 - d. shown only if the lipid is in a nonpolar solvent.
 - e. important in energy storage by lipids.
10. Which of the following statements about condensation reactions is *not* true?
 - a. Protein synthesis results from them.
 - b. Polysaccharide synthesis results from them.
 - c. They involve covalent bonds.
 - d. They consume water as a reactant.
 - e. Different condensation reactions produce different kinds of macromolecules.

FOR DISCUSSION

1. Suppose that, in a given protein, one lysine is replaced by aspartic acid (see Table 3.1). Does this change occur in the primary structure or in the secondary structure? How might it result in a change in tertiary structure? In quaternary structure?
2. If there are 20 different amino acids commonly found in proteins, how many different dipeptides are there? How many different tripeptides?

ADDITIONAL INVESTIGATION

Human hair is composed of a protein, keratin. At the hair salon, two techniques are used to modify the three-dimensional shape of hair. Styling involves heat, and a perm involves cleaving and

reforming disulfide bonds. How would you investigate these phenomena in terms of protein structure?

WORKING WITH DATA (GO TO yourBioPortal.com)

Primary Structure Specifies Tertiary Structure In this hands-on exercise based on Figure 3.9, you will learn about the methods used to disrupt the chemical interactions that determine

the tertiary structure of proteins. You will examine the original data that led Anfinsen to conclude that denaturation of ribonuclease is reversible.

4

Nucleic Acids and the Origin of Life

Looking for life

The trip had lasted a long and anxious ten months when, in the summer of 1976, the first of two visitors from Earth landed on a plain on the Martian surface. A second spacecraft arrived in September. The task of these robotic laboratories, part of NASA's Viking project, was to search for life.

On Earth, life has existed for several billion years and has spread over most of the planet's surface. Determining life's origins is difficult, however, because (with few exceptions) simple organisms leave no fossils. On Mars, scientists thought, things might be different. A primitive form of life might exist there now, or might have left chemical signatures that remain in place, untouched by other organisms.

The two Viking spacecraft that landed on Mars in 1976 analyzed soil samples for the small molecules of life, in-

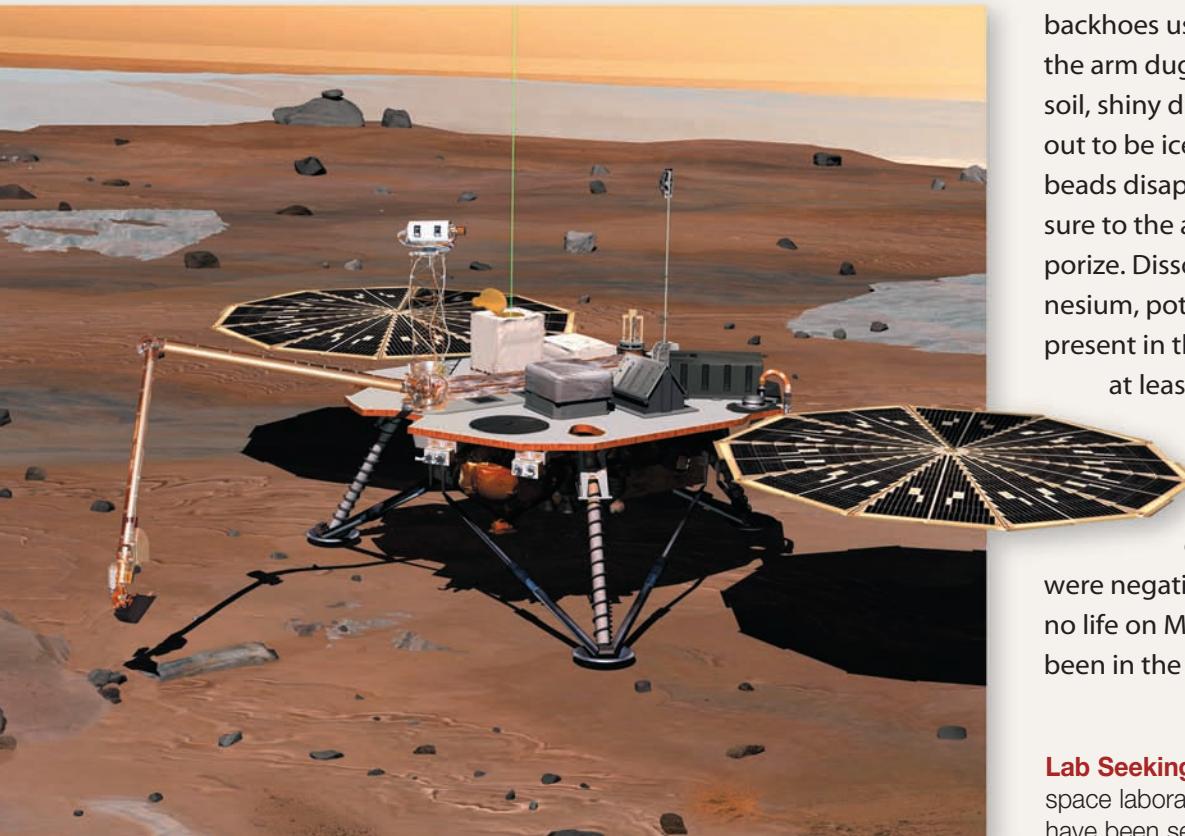
cluding simple sugars and amino acids. None were found. The robotic laboratories immersed soil samples in an aqueous solution of sugars, amino acids, and minerals. Living organisms take in and break down such substances from their environment, releasing gases such as CO₂. A small amount of CO₂ was detected in one experiment, but, frustratingly, no gases were released in further experiments.

The results from the Viking landers remain controversial. Why did that one experiment detect a sign of life? The 1976 robotic landers are still on Mars but have long since stopped working. In 2008, more probes were sent from Earth, carrying more sophisticated instruments. One of them, the Phoenix lander, is in a northern region of Mars, at a latitude corresponding to that of Alaska on

Earth. Phoenix has a robotic arm like the backhoes used in a construction site. When the arm dug a small trench into the Martian soil, shiny dice-sized beads of what turned out to be ice were exposed, although the beads disappeared in a few days as exposure to the atmosphere caused them to vaporize. Dissolved ions such as sodium, magnesium, potassium and chloride were all present in the frozen water, indicating that

at least those requirements for life are present on Mars. Once again, the soil was analyzed for traces of current or past organisms; once again, the results were negative. But even if there probably is no life on Mars today, there might have been in the past.

Lab Seeking Life Landers such as the robotic space laboratory Phoenix, shown here on Earth, have been sent to look for traces of life on Mars.





Ice on Mars The Phoenix landing site (blue dot) is near the Martian north pole, where chemical traces of life might be preserved in the hypercold environment. When the lander scooped up a patch of soil for analysis, it also took photos that revealed ice crystals just below the surface of the Red Planet.

As we saw in Chapter 2, water is a key requirement for life. Remote measurements from orbiting spacecraft and chemical measurements using special telescopes have shown that water is present on Mars and, indeed, on some of the moons of other planets in our solar system.

Scientists are using their knowledge of the small and large molecules that are present in living organisms to search for the chemical signatures of life on other planets. Chapters 2 and 3 described molecules that are important for biological structure and function. In Chapter 4, we turn to certain molecules involved in the origin and perpetuation of life itself.

IN THIS CHAPTER we first describe the structure of nucleic acids, the informational macromolecules needed for the perpetuation of life. We then turn to biologists' speculations on the origin of life and describe early experimental evidence that life on Earth today comes from pre-existing life. We present some ideas on the formation of the building blocks of life, including the monomers and polymers that characterize biological systems. Finally, we describe some proposals for the origin of cells.

CHAPTER OUTLINE

- 4.1** What Are the Chemical Structures and Functions of Nucleic Acids?
- 4.2** How and Where Did the Small Molecules of Life Originate?
- 4.3** How Did the Large Molecules of Life Originate?
- 4.4** How Did the First Cells Originate?

4.1 What Are the Chemical Structures and Functions of Nucleic Acids?

From medicine to evolution, from agriculture to forensics, the properties of nucleic acids impact our lives every day. It is with nucleic acids that the concept of “information” entered the biological vocabulary. Nucleic acids are uniquely capable of coding for and transmitting biological information.

The **nucleic acids** are polymers specialized for the storage, transmission between generations, and use of genetic information. There are two types of nucleic acids: **DNA** (*deoxyribonucleic acid*) and **RNA** (*ribonucleic acid*). DNA is a macromolecule that encodes hereditary information and passes it from generation to generation. Through an RNA intermediate, the information encoded in DNA is used to specify the amino acid sequences of proteins. Information flows from DNA to DNA during reproduction. In the non-reproductive activities of the cell, information flows from DNA to RNA to proteins. It is the proteins that ultimately carry out life’s functions.

Nucleotides are the building blocks of nucleic acids

Nucleic acids are composed of monomers called **nucleotides**, each of which consists of a pentose sugar, a phosphate group, and a nitrogen-containing **base**. (Molecules consisting of a pentose sugar and a nitrogenous base—but no phosphate group—are called *nucleosides*.) The bases of the nucleic acids take one of two chemical forms: a six-membered single-ring structure called a **pyrimidine**, or a fused double-ring structure called a **purine** (Figure 4.1). In DNA, the pentose sugar is **deoxyribose**, which differs from the **ribose** found in RNA by the absence of one oxygen atom (see Figure 3.14).

In both RNA and DNA, the backbone of the macromolecule consists of a chain of alternating pentose sugars and phosphate groups (sugar–phosphate–sugar–phosphate). The bases are attached to the sugars and project from the polynucleotide chain (Figure 4.2). The nucleotides are joined by **phosphodiester linkages** between the sugar of one nucleotide and the phosphate of the next (*diester* refers to the two covalent bonds formed by —OH groups reacting with acidic phosphate groups). The phosphate groups link carbon 3 in one pentose sugar to carbon 5 in the adjacent sugar.

Most RNA molecules consist of only one polynucleotide chain. DNA, however, is usually double-stranded; its two polynucleotide chains are held together by hydrogen bonding between their nitrogenous bases. The two strands of DNA run in opposite directions. You can see what this means by drawing an arrow through a phosphate group from carbon 5 to

4.1 Nucleotides Have Three Components Nucleotide monomers are the building blocks of DNA and RNA polymers.

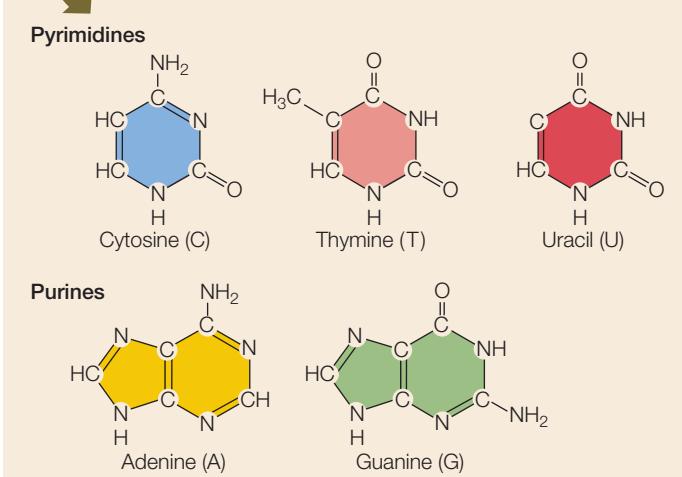
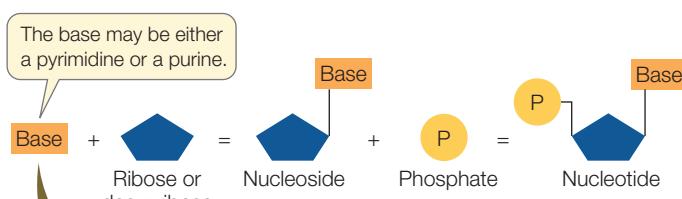
yourBioPortal.com

GO TO Web Activity 4.1 • Nucleic Acid Building Blocks

carbon 3 in the next ribose. If you do this for both strands of the DNA in Figure 4.2, the arrows will point in opposite directions. This *antiparallel* orientation allows the strands to fit together in three-dimensional space.

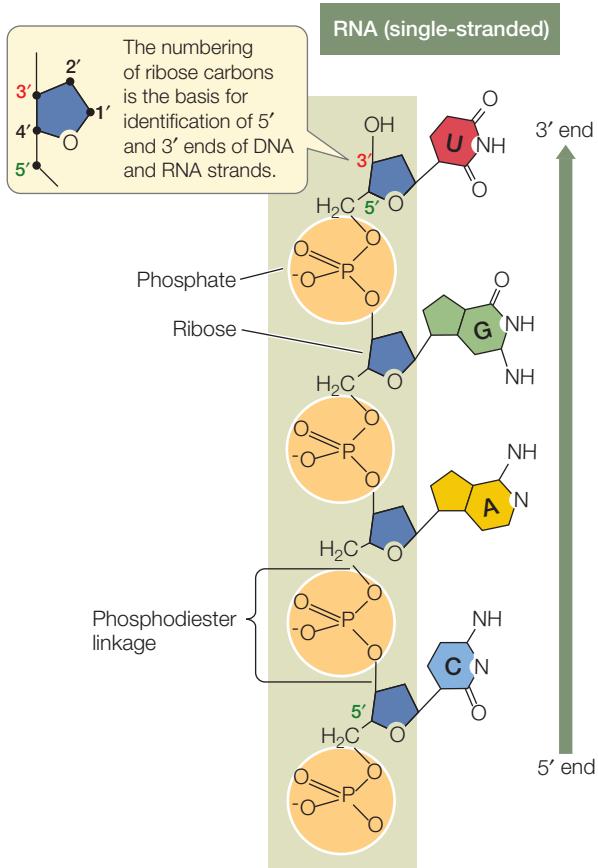
Base pairing occurs in both DNA and RNA

Only four nitrogenous bases—and thus only four nucleotides—are found in DNA. The DNA bases and their abbreviations are **adenine (A)**, **cytosine (C)**, **guanine (G)**, and **thymine (T)**. Adenine and guanine are purines; thymine and cytosine are pyrimidines. RNA is also made up of four different monomers, but its nucleotides differ from those of DNA. In RNA the nucleotides are termed *ribonucleotides* (the ones in DNA are *deoxyribonucleotides*). They contain ribose rather than deoxyribose, and in-

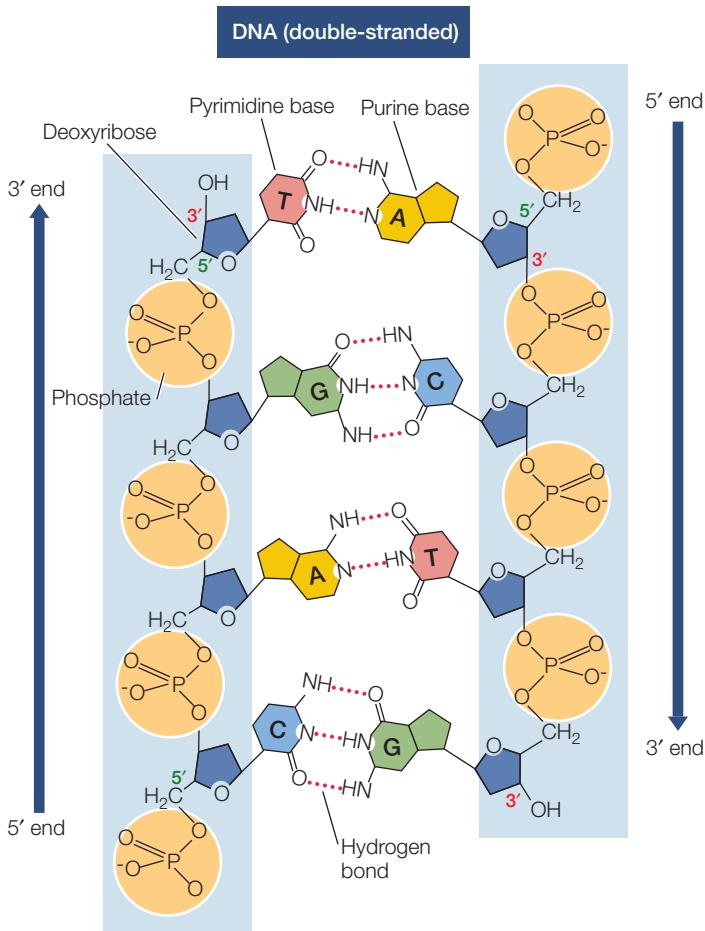


4.2 Distinguishing Characteristics of DNA and RNA Polymers

RNA is usually a single strand. DNA usually consists of two strands running in opposite directions (antiparallel).



In RNA, the bases are attached to ribose. The bases in RNA are the purines adenine (A) and guanine (G) and the pyrimidines cytosine (C) and uracil (U).



In DNA, the bases are attached to deoxyribose, and the base thymine (T) is found instead of uracil. Hydrogen bonds between purines and pyrimidines hold the two strands of DNA together.

TABLE 4.1
Distinguishing RNA from DNA

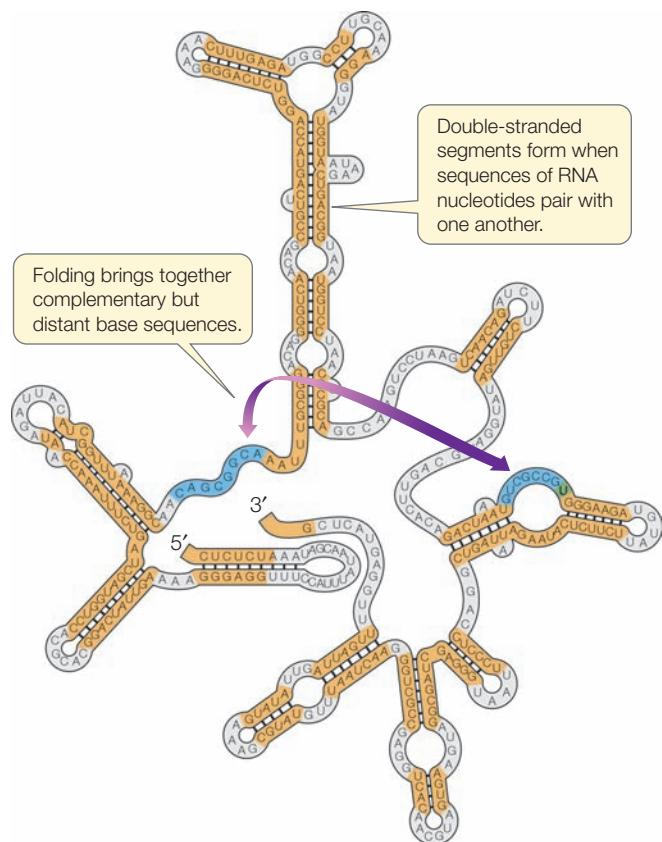
NUCLEIC ACID	SUGAR	BASES	STRANDS
RNA	Ribose	Adenine	Single
		Cytosine	
		Guanine	
		Uracil	
DNA	Deoxyribose	Adenine	Double
		Cytosine	
		Guanine	
		Thymine	

stead of the base thymine, RNA uses the base **uracil (U)**. The other three bases are the same in RNA and DNA (**Table 4.1**).

The key to understanding the structure and function of nucleic acids is the principle of **complementary base pairing**. In double-stranded DNA, adenine and thymine always pair (A-T), and cytosine and guanine always pair (C-G).

Three factors make base pairing complementary:

- The sites for hydrogen bonding on each base
- The geometry of the sugar–phosphate backbone, which brings complementary bases near each other



4.3 Hydrogen Bonding in RNA When a single-stranded RNA folds on itself, hydrogen bonds between complementary sequences can stabilize it into a three-dimensional shape with complicated surface characteristics.

- The molecular sizes of the paired bases; the pairing of a larger purine with a smaller pyrimidine ensures stability and uniformity in the double-stranded molecule of DNA

Although RNA is generally single-stranded, complementary hydrogen bonding between ribonucleotides plays important roles in determining the three-dimensional shapes of some types of RNA molecules, since portions of the single-stranded RNA can fold back and pair with each other (**Figure 4.3**). Complementary base pairing can also take place between ribonucleotides and deoxyribonucleotides. In RNA, guanine and cytosine pair (G-C), as in DNA, but adenine pairs with uracil (A-U). Adenine in an RNA strand can pair either with uracil (in another RNA strand) or with thymine (in a DNA strand).

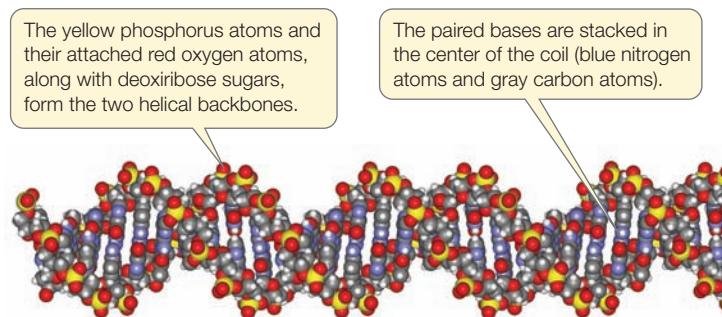
The three-dimensional physical appearance of DNA is strikingly uniform. The segment shown in **Figure 4.4** could be from any DNA molecule. The variations in DNA—the different sequences of bases—are strictly internal. Through hydrogen bonding, the two complementary polynucleotide strands pair and twist to form a **double helix**. When compared with the complex and varied tertiary structures of proteins, this uniformity is surprising. But this structural contrast makes sense in terms of the functions of these two classes of macromolecules. As we describe in Section 3.2, the different and unique shapes of proteins permit these macromolecules to recognize specific “target” molecules. The area on the surface of a protein that interacts with the target molecule must match the shape of at least part of the target molecule. In other words, structural diversity in the target molecules requires corresponding diversity in the structures of the proteins themselves. Structural diversity is necessary in DNA as well. However, the diversity of DNA is found in its base sequence rather than in the physical shape of the molecule. Different DNA base sequences encode specific information.

yourBioPortal.com

GO TO Web Activity 4.2 • DNA Structure

DNA carries information and is expressed through RNA

DNA is a purely *informational* molecule. The information is encoded in the sequence of bases carried in its strands—the infor-



4.4 The Double Helix of DNA The backbones of the two strands in a DNA molecule are coiled in a double helix that is held together by hydrogen bonds between the purines and pyrimidines in the interior of the structure. In this model, the small white atoms represent hydrogen.

4.5 DNA Stores Information The DNA macromolecule stores information that can either be copied (replicated) or transcribed into RNA. RNA can then be translated into protein.

mation encoded in the sequence TCAGCA is different from the information in the sequence CCAGCA. DNA has two functions in terms of information. Taken together, they comprise the *central dogma of molecular biology* (Figure 4.5).

- DNA can reproduce itself exactly. This is called *DNA replication*. It is done by polymerization on a template.
- DNA can copy its information into RNA, in a process called *transcription*. The nucleotide sequence in RNA can specify a sequence of amino acids in a polypeptide. This is called *translation*.

While the details of these important processes are described in later chapters, it is important to realize two things at this point:

1. *DNA replication and transcription depend on the base pairing properties of nucleic acids.* The hydrogen-bonded base pairs are A-T and G-C in DNA and A-U and G-C in RNA (see Figure 4.2). Consider this double-stranded DNA region:

```
T CAGCA
      AGTCGT
```

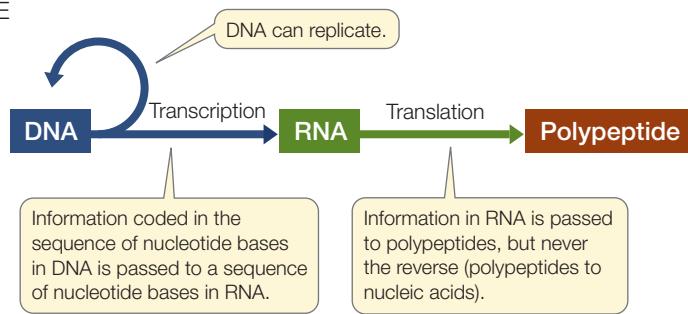
Transcription of the lower strand will result in a single strand of RNA with the sequence UCAGCA. Can you figure out what the top strand would produce?

2. *DNA replication usually involves the entire DNA molecule, but only relatively small sections of the DNA are transcribed into RNA molecules.* Since DNA holds essential information, it must be replicated completely so that each new cell or new organism receives a complete set of DNA from its parent. The complete set of DNA in a living organism is called its **genome**. However, not all of the information in the genome is needed at all times (Figure 4.6A).

The sequences of DNA that encode specific proteins are transcribed into RNA and are called **genes** (Figure 4.6B). In humans, the genes that encode the subunits of the protein hemoglobin (see Figure 3.10) are expressed only in the precursors of red blood cells. The genetic information in each globin gene is transcribed into RNA and then translated into a globin polypeptide. In other tissues, such as the muscles, the genes that encode the globin subunits are not transcribed, but others are—for example, the genes for the myosin proteins that are the major component of muscle fibers (see Section 48.1).

The DNA base sequence reveals evolutionary relationships

Because DNA carries hereditary information from one generation to the next, a theoretical series of DNA molecules, with changes in base sequences, stretches back through the lineage of every organism to the beginning of biological evolution on Earth, about 4 billion years ago. Therefore, closely related living species should have more similar base sequences than species that are more distantly related. The details of how scientists use this information are covered in Chapter 24.

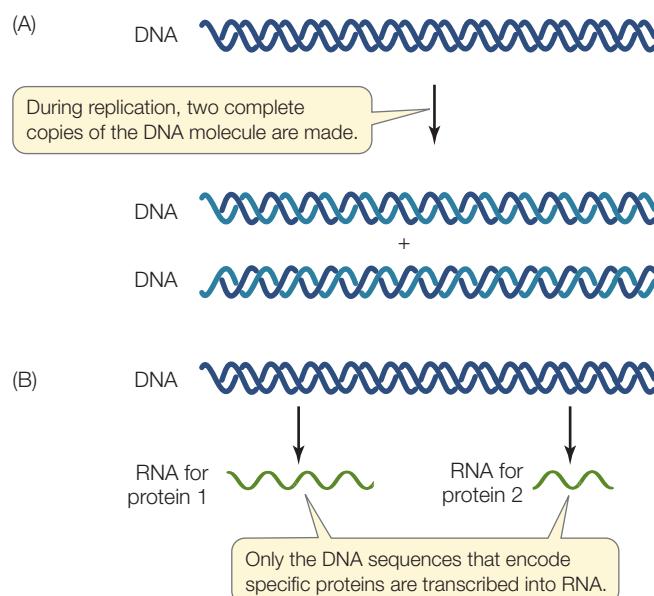


The elucidation and examination of DNA base sequences has confirmed many of the evolutionary relationships that were inferred from more traditional comparisons of body structures, biochemistry, and physiology. Many studies of anatomy, physiology, and behavior have concluded that the closest living relative of humans (*Homo sapiens*) is the chimpanzee (genus *Pan*). In fact, the chimpanzee genome shares more than 98 percent of its DNA base sequence with the human genome. Increasingly, scientists turn to DNA analyses to elucidate evolutionary relationships when other comparisons are not possible or are not conclusive. For example, DNA studies revealed a close relationship between starlings and mockingbirds that was not expected on the basis of their anatomy or behavior.

Nucleotides have other important roles

Nucleotides are more than just the building blocks of nucleic acids. As we will describe in later chapters, there are several nucleotides with other functions:

- ATP (adenosine triphosphate) acts as an energy transducer in many biochemical reactions (see Section 8.2).



4.6 DNA Replication and Transcription DNA is usually completely replicated (A) but only partially transcribed (B). RNA transcripts encode the genes for specific proteins. Transcription of the many different proteins is activated at different times and, in multicellular organisms, in different cells of the body.

- GTP (guanosine triphosphate) serves as an energy source, especially in protein synthesis. It also plays a role in the transfer of information from the environment to cells (see Section 7.2).
- cAMP (cyclic adenosine monophosphate) is a special nucleotide in which an additional bond forms between the sugar and phosphate group. It is essential in many processes, including the actions of hormones and the transmission of information by the nervous system (see Section 7.3).

4.1 RECAP

The nucleic acids DNA and RNA are polymers made up of nucleotide monomers. The sequence of nucleotides in DNA carries the information that is used by RNA to specify primary protein structure. The genetic information in DNA is passed from generation to generation and can be used to understand evolutionary relationships.

- List the key differences between DNA and RNA. Between purines and pyrimidines. **See p. 61, Table 4.1, and Figure 4.1**
- How do purines and pyrimidines pair up in complementary base pairing? **See pp. 62–63 and Figure 4.2**
- What are the differences between DNA replication and transcription? **See pp. 63–64 and Figures 4.5 and 4.6**
- How can DNA molecules be very diverse, even though they appear to be structurally similar? **See p. 64**

We have seen that the nucleic acids RNA and DNA carry the blueprint of life, and that the inheritance of these macromolecules reaches back to the beginning of evolutionary time. But when, where, and how did nucleic acids arise on Earth? How did the building blocks of life such as amino acids and sugars originally arise?

4.2 How and Where Did the Small Molecules of Life Originate?

Chapter 2 points out that living things are composed of the same atomic elements as the inanimate universe—the 92 naturally occurring elements of the periodic table (see Figure 2.2). But the arrangements of these atoms into molecules are unique in biological systems. You will not find biological molecules in inanimate matter unless they came from a once-living organism.

It is impossible to know for certain how life on Earth began. But one thing is sure: life (or at least life as we know it) is not constantly being re-started. That is, *spontaneous generation* of life from inanimate nature is not happening before our eyes. Now and for many millenia past, all life has come from life that existed before. But people, including scientists, did not always believe this.

Experiments disproved spontaneous generation of life

The idea that life could have originated from nonliving matter is common in many cultures and religions. During the European Renaissance (from about 1450 to 1700, a period that witnessed the birth of modern science), most people thought that at least some forms of life arose repeatedly and directly from inanimate or decaying matter by *spontaneous generation*. Many thought that mice arose from sweaty clothes placed in dim light; that frogs sprang directly from moist soil; and that rotting meat produced flies. Scientists such as the Italian physician and poet Francesco Redi, however, doubted these assumptions. Redi proposed that flies arose not by some mysterious transformation of decaying meat, but from other flies that laid their eggs on the meat. In 1668, Redi performed a scientific experiment—a relatively new concept at the time—to test his hypothesis. He set out several jars containing chunks of meat.

- One jar contained meat exposed to both air and flies.
- A second jar was covered with a fine cloth so that the meat was exposed to air, but not to flies.
- The third jar was sealed so the meat was exposed to neither air nor flies.

As he had hypothesized, Redi found maggots, which then hatched into flies, only in the first jar. This finding demonstrated that maggots could occur only where flies were present. The idea that a complex organism like a fly could appear *de novo* from a nonliving substance in the meat, or from “something in the air,” was laid to rest. Well, perhaps not quite to rest.

In the 1660s, newly developed microscopes revealed a vast new biological world. Under microscopic observation, virtually every environment on Earth was found to be teeming with tiny organisms. Some scientists believed these organisms arose spontaneously from their rich chemical environment, by the action of a “life force.” But experiments by the great French scientist Louis Pasteur showed that microorganisms can arise only from other microorganisms, and that an environment without life remains lifeless (**Figure 4.7**).

yourBioPortal.com

GO TO [Animated Tutorial 4.1 • Pasteur's Experiment](#)

Pasteur’s and Redi’s experiments showed that living organisms cannot arise from nonliving materials *under the conditions that existed on Earth during their lifetimes*. But their experiments did not prove that spontaneous generation never occurred. Eons ago, conditions on Earth and in the atmosphere above it were vastly different. Indeed, conditions similar to those found on primitive Earth may have existed, or may exist now, on other bodies in our solar system and elsewhere. This has led scientists to ask whether life has originated on other bodies in space, as it did on Earth.

Life began in water

As we emphasize in Chapter 2 and in the opening story of this chapter, the presence of water on a planet or moon is a necessary prerequisite for life as we know it. Astronomers believe our solar system began forming about 4.6 billion years ago, when a

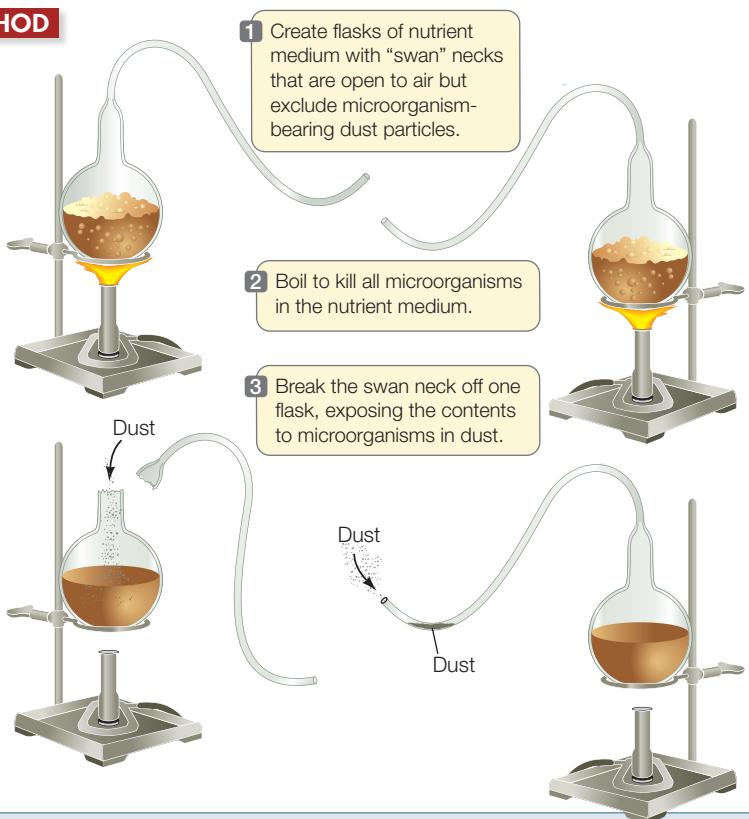
INVESTIGATING LIFE

4.7 Disproving the Spontaneous Generation of Life

Previous experiments disproving spontaneous generation were called into question in regard to microorganisms, whose abundance and diversity were appreciated but whose living processes were not understood. Louis Pasteur's classic experiments disproved the spontaneous generation of microorganisms.

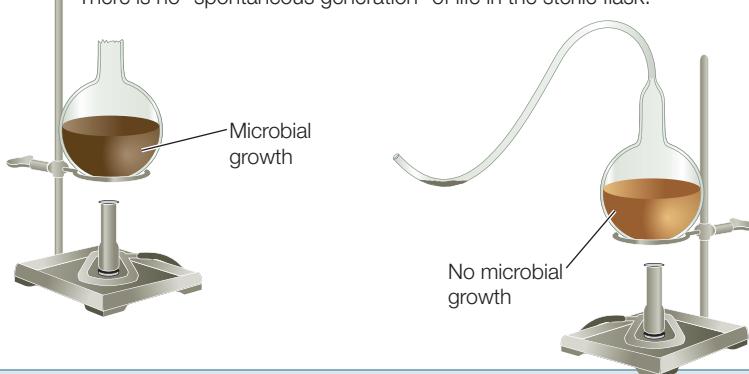
HYPOTHESIS Microorganisms come only from other microorganisms and cannot arise by spontaneous generation.

METHOD



RESULTS

Microbial life grows only in the flasks exposed to microorganisms. There is no "spontaneous generation" of life in the sterile flask.



CONCLUSION All life comes from pre-existing life. An environment without life remains lifeless.

Go to yourBioPortal.com for original citations, discussions, and relevant links for all INVESTIGATING LIFE figures.

star exploded and collapsed to form the sun and 500 or so bodies, called planetesimals. These planetesimals collided with one another to form the inner planets, including Earth and Mars. The first chemical signatures indicating the presence of life on Earth appear to be about 4 billion years old. So it took 600 million years, during a geological time frame called the Hadean, for the chemical conditions on Earth to become just right for life. Key among those conditions was the presence of water.

Ancient Earth probably had a lot of water high in its atmosphere. But the new planet was hot, and the water remained in vapor form and dissipated into space. As Earth cooled, it became possible for water to condense on the planet's surface—but where did that water come from? One current view is that comets (loose agglomerations of dust and ice that have orbited the sun since the planets formed) struck Earth and Mars repeatedly, bringing to those planets not only water but other chemical components of life, such as nitrogen.

As the planets cooled and chemicals from their crusts dissolved in the water, simple chemical reactions would have taken place. Some of these reactions might have led to life, but impacts by large comets and rocky meteorites released enough energy to heat the developing oceans almost to boiling, thus destroying any early life. On Earth, these large impacts eventually subsided, and some time around 3.8 to 4 billion years ago life gained a foothold. There has been life on Earth ever since. Because Mars and some other celestial bodies have a similar geological history, the possibility exists that life exists or has existed on them. This possibility was an impetus for sending the Viking and Phoenix landers to Mars.

Several models have been proposed to explain the origin of life on Earth. The next sections discuss two alternative theories: that life came from outside of Earth, or that life arose on Earth through chemical evolution.

Life may have come from outside Earth

In 1969, a remarkable event led to the discovery that a meteorite from space carried molecules that were characteristic of life on Earth. On September 28 of that year, fragments of a meteorite fell around the town of Murchison, Australia. Using gloves to avoid Earth-derived contamination, scientists immediately shaved off tiny pieces of the rock, put them in



4.8 The Murchison Meteorite Pieces from a fragment of the meteorite that landed in Australia in 1969 were put into test tubes with water. Soluble molecules present in the rock, including amino acids, nucleotide bases, and sugars, dissolved in the water. Plastic gloves and sterile instruments were used to reduce the possibility of contamination with substances from Earth.

test tubes and extracted them in water (**Figure 4.8**). They found a number of the molecules that are unique to life, including purines, pyrimidines, sugars, and ten amino acids.

Were these molecules truly brought from space as part of the meteorite, or did they get there after the rock landed on Earth? There were a number of reasons to believe the molecules were not Earthly contaminants:

- The scientists took great care to avoid contamination. They used gloves and sterile instruments, took pieces from below the rock's surface, and did their work very soon after it landed (hopefully before Earth organisms could contaminate the samples).
- Amino acids found in living organisms on Earth are left-handed (see Figure 3.2). The amino acids in the meteorite were a mixture of right- and left-handed forms, with a slight preponderance of the left-handed. Thus the amino acids in the meteorite were not likely to have come from a living organism on Earth.
- In the story that opens Chapter 2, we describe how the ratio of isotopes in a living organism reflects that isotope ratio in the environment where the organism lives. The isotope ratios for carbon and hydrogen in the sugars from the meteorite were different from the ratios of those elements found on Earth.

In 1984, another informative meteorite, this one the size of a softball, was found in Antarctica. We know that the meteorite, ALH 84001, came from Mars because the composition of the gases trapped within the rock was identical to the composition found in the Martian atmosphere, which is quite different from Earth's atmosphere. Radioactive dating and mineral analyses determined that ALH 84001 was 4.5 billion years old and was blasted off the Martian surface 16 million years ago. It landed on Earth fairly recently, about 13,000 years ago.

Scientists found water trapped below the Martian meteorite's surface. This discovery was not surprising, given that surface observations had already shown that water was once abundant on Mars (see the chapter-opening story). Because water is essential for life, scientists wondered whether the meteorite might contain other signs of life as well. Their analysis revealed two substances related to living systems. First, simple carbon-containing molecules called polycyclic aromatic hydrocarbons were present in small but unmistakable amounts; these substances can be formed by living organisms. Second, crystals of magnetite, an iron oxide mineral made by many living organisms on Earth, were found in the interior of the rock.

ALH 84001 and the Murchison meteorite are not the only visitors from outer space that have been shown to contain chemical signatures of life. While the presence of such molecules in rocks may suggest that those rocks once harbored life, it does not prove that there were living organisms in the rocks when they landed on Earth. Most scientists find it hard to believe that an organism could survive thousands of years of traveling through space in a meteorite, followed by intense heat as the meteorite passed through Earth's atmosphere. But there is some evidence that the heat inside some meteorites may not have been severe. When weakly magnetized rock is heated, it reorients its magnetic field to align with the magnetic field around it. In the case of ALH 84001, this would have been Earth's powerful magnetic field, which would have affected the meteorite as it approached our planet.

Careful measurements indicate that, while reorientation did occur at the surface of the rock, it did not occur on the inside. The scientists who took these measurements, Benjamin Weiss and Joseph Kirschvink at the California Institute of Technology, concluded that the inside of ALH 84001 was never heated over 40°C as it entered Earth's atmosphere. This suggests that a long interplanetary trip by living organisms could be possible.

Prebiotic synthesis experiments model the early Earth

It is clear that other bodies in the solar system have, or once had, water and other simple molecules. Possibly, a meteorite was the source of the simple molecules that were the original building blocks for life on Earth. But a second theory for the origin of life on Earth, **chemical evolution**, holds that conditions on primitive Earth led to the formation of these simple molecules (prebiotic synthesis), and these molecules led to the formation of life forms. Scientists have sought to reconstruct those primitive conditions, both physically (hot or cold) and chemically (by re-creating the combinations and proportions of elements that may have been present).

HOT CHEMISTRY The amounts of trace metals such as molybdenum and rhenium in sediments under oceans and lakes is directly proportional to the amount of oxygen gas (O₂) present in and above the water. Measurements of dated sedimentary cores indicate that none of these rare metals was present prior to 2.5 billion years ago. This and other lines of evidence suggest that there was little oxygen gas in Earth's early atmosphere. Oxygen gas is thought to have accumulated about 2.5 billion years ago.

as the by-product of photosynthesis by single-celled life forms; today 21 percent of our atmosphere is O₂.

In the 1950s, Stanley Miller and Harold Urey at the University of Chicago set up an experimental “atmosphere” containing the gases thought to have been present in Earth’s early atmosphere: hydrogen gas, ammonia, methane gas, and water vapor. They passed an electric spark through these gases, to simulate lightning as a source of energy to drive chemical reactions. Then, they cooled the system so the gases would condense and collect in a watery solution, or “ocean” (Figure 4.9). After a few days of continuous operation, the system contained numerous complex molecules, including amino acids, purines, and pyrimidines—some of the building blocks of life.

yourBioPortal.com

GO TO Animated Tutorial 4.2 • Synthesis of Prebiotic Molecules

The results of this experiment were profoundly important in giving weight to speculations about the chemical origin of life on Earth and elsewhere in the universe. Decades of experimental work and critical evaluation followed. The experiments showed that, under the conditions used by Miller and Urey, many small molecular building blocks of life could be formed:

- All five bases that are present in DNA and RNA (i.e., A, T, C, G and U)
- 17 of the 20 amino acids used in protein synthesis
- 3- to 6-carbon sugars

However, the 5-carbon sugar ribose was not produced in these experiments.

In science, an experiment and its results must be repeated, reinterpreted, and refined as more knowledge accumulates. The results of the Miller–Urey experiments have undergone several such refinements.

The amino acids in living things are always L-isomers (see Figure 3.2 and p. 43). But a mixture of D- and L-isomers appeared in the amino acids formed in the Miller–Urey experiments. Recent experiments show that natural processes could have selected the L-amino acids from the mixture. Some minerals, especially calcite-based rocks, have unique crystal structures that selectively bind to D- or L-amino acids, separating the two. Such rocks were abundant on early Earth. This suggests that while both kinds of amino acid structures were made, binding to certain rocks may have eliminated the D- amino

acids. (Interestingly, some meteorites, such as the Murchison meteorite, also have this selectivity.)

Ideas about Earth’s original atmosphere have changed since Miller and Urey did their experiments. There is abundant evidence indicating that major volcanic eruptions occurred 4 bil-

INVESTIGATING LIFE

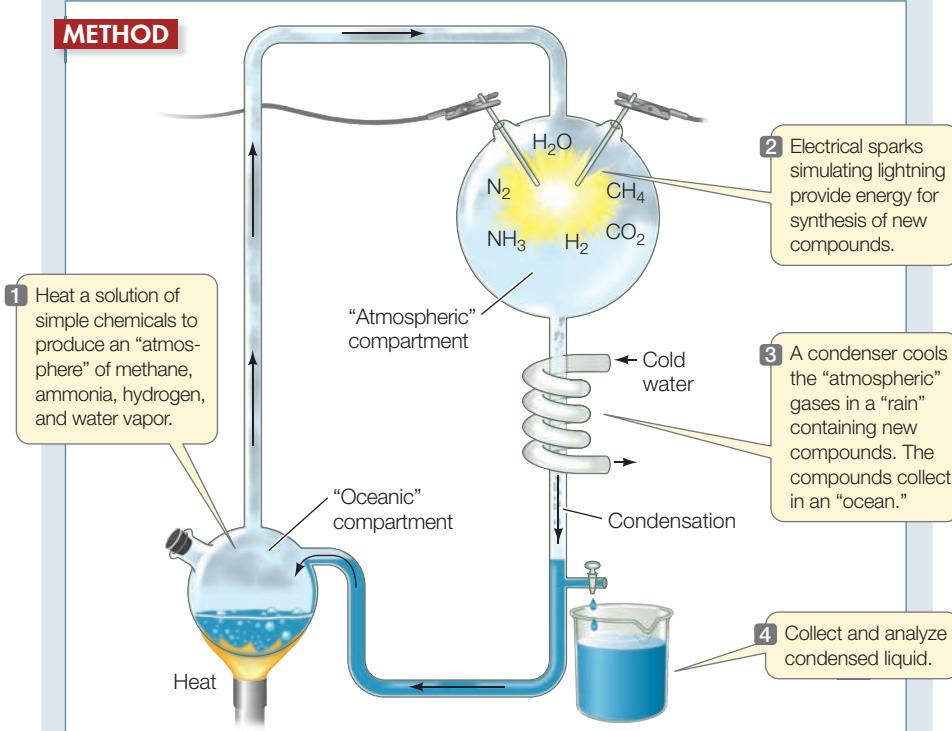
4.9 Miller and Urey Synthesized Prebiotic Molecules in an Experimental Atmosphere

With an increased understanding of the atmospheric conditions that existed on primitive Earth, the researchers devised an experiment to see if these conditions could lead to the formation of organic molecules.

HYPOTHESIS

Organic chemical compounds can be generated under conditions similar to those that existed in the atmosphere of primitive Earth.

METHOD



RESULTS



Reactions in the condensed liquid eventually formed organic chemical compounds, including purines, pyrimidines, and amino acids.

CONCLUSION

The chemical building blocks of life could have been generated in the probable atmosphere of early Earth.

FURTHER INVESTIGATION: What result would you predict if O₂ were present in the “atmosphere” in this experiment?

Go to **yourBioPortal.com** for original citations, discussions, and relevant links for all INVESTIGATING LIFE figures.

lion years ago, which would have released carbon dioxide (CO_2), nitrogen (N_2), hydrogen sulfide (H_2S), and sulfur dioxide (SO_2) into the atmosphere. Experiments using these gases in addition to the ones in the original experiment have produced more diverse molecules, including:

- Vitamin B₆, pantothenic acid (a component of coenzyme A), and nicotinamide (part of NAD, which is involved in energy metabolism).
- Carboxylic acids such as succinic and lactic acids (also involved in energy metabolism) and fatty acids.
- Ribose, a key component of RNA, which can be formed from formaldehyde gas (HCHO), evidence of which has been found in space.

COLD CHEMISTRY Stanley Miller also performed a long-term experiment in which the electric spark was not used. In 1972, he filled test tubes with ammonia gas, water vapor and cyanide (HCN), another molecule that is thought to have formed on primitive Earth. After checking that there were no contaminating substances or organisms that might confound the results, he sealed the tubes and cooled them to -100°C , the temperature of the ice that covers Europa, one of Jupiter's moons. Opening the tubes 25 years later, he found amino acids and nucleotide bases. Apparently, pockets of liquid water within the ice had allowed high concentrations of the starting materials to accumulate, thereby speeding up chemical reactions. The important conclusion is that the cold water within ice on ancient Earth, and other celestial bodies such as Mars, Europa, and Enceladus (one of Saturn's moons; satellite photos have revealed geysers of liquid water coming from its interior) may have provided environments for the prebiotic synthesis of molecules required for the subsequent formation of simple living systems.

4.2 RECAP

Life does not arise repeatedly through spontaneous generation, but comes from pre-existing life. Water is an essential ingredient for the emergence of life. Meteorites that have landed on Earth provide some evidence for an extraterrestrial origin of life. Prebiotic chemical synthesis experiments provide support for the idea that life's simple molecules formed in the primitive Earth environment.

- Explain how Redi's and Pasteur's experiments disproved spontaneous generation. **See p. 65 and Figure 4.7**
- What is the evidence that life on Earth came from other bodies in the solar system? **See pp. 66–67**
- What is the significance of the Miller–Urey experiment, what did it find, and what were its limitations? **See p. 68 and Figure 4.9**

Chemistry experiments using conditions modeling the ancient Earth's environment suggest an origin for the monomers (such as amino acids) that make up the polymers (such as proteins)

that characterize life. How did these polymers develop on the ancient Earth?

4.3 How Did the Large Molecules of Life Originate?

The Miller–Urey experiment and other experiments that followed it provide a plausible scenario for the formation of the building blocks of life under conditions that prevailed on primitive Earth. The next step in forming and supporting a general theory on the origin of life on Earth would be an explanation of the formation of polymers from these monomers.

Chemical evolution may have led to polymerization

Scientists have used a number of model systems to try to simulate conditions under which polymers might have been made. Each of these systems is based on several observations and speculations:

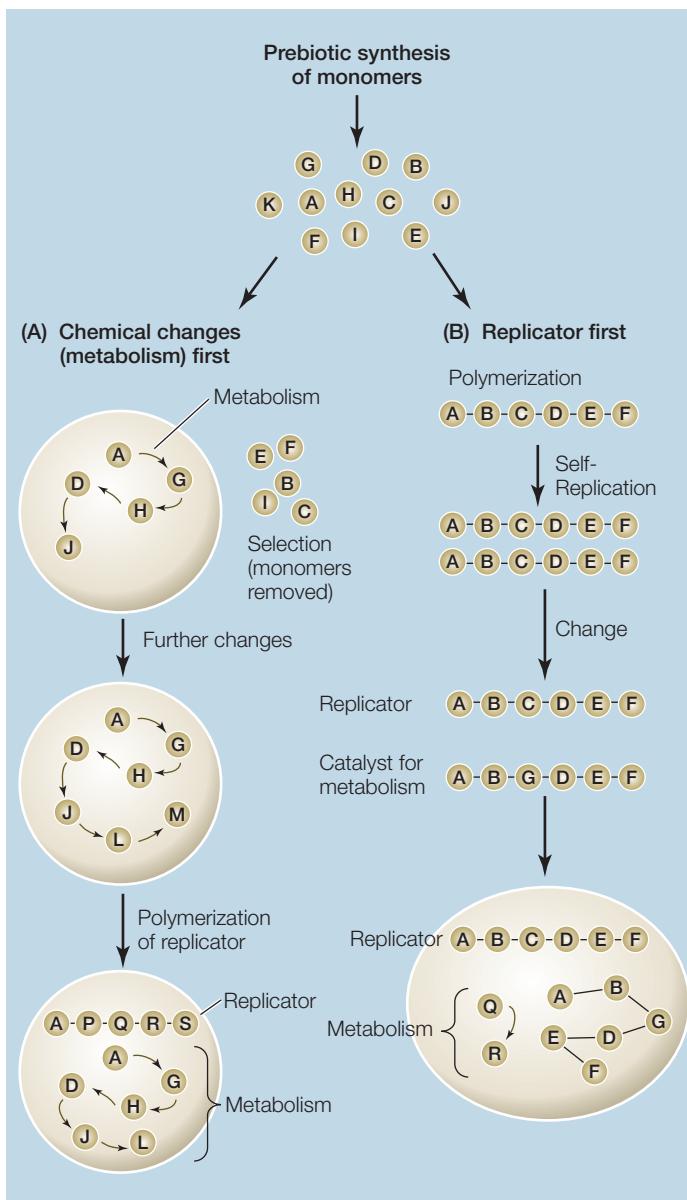
- *Solid mineral surfaces*, such as powder-like clays, have large surface areas. Scientists speculate that the silicates within clay may have been catalytic (speeded up the reactions) in the formation of early carbon-based molecules.
- *Hydrothermal vents* deep in the ocean, where hot water emerges from beneath Earth's crust, lack oxygen gas and contain metals such as iron and nickel. In laboratory experiments, these metals have been shown to catalyze the polymerization of amino acids to polypeptides in the absence of oxygen.
- *Hot pools* at the edges of oceans may, through evaporation, have concentrated monomers to the point where polymerization was favored (the "primordial soup" hypothesis).

In whatever ways the earliest stages of chemical evolution occurred, they resulted in the emergence of monomers and polymers that have probably remained unchanged in their general structure and function for several billion years.

There are two theories for the emergence of nucleic acids, proteins, and complex chemistry

Earlier in this chapter, we described the key roles of nucleic acids as informational molecules that are passed on from one generation to the next. We also described how DNA is transcribed to RNA, which can then be translated into protein (see Figure 4.5). Chapter 3 describes the roles of proteins as catalysts, speeding up biochemical transformations (see Section 3.2). In existing life forms, nucleic acids and proteins require one another in order to perpetuate life. For the origin of life, this results in a chicken-or-egg problem. Which came first, the genetic material (nucleic acids) or proteins? Two ideas have emerged. One suggests that sequential catalytic changes (primitive metabolism) came first. The other suggests that replication by nucleic acids preceded metabolism (Figure 4.10).

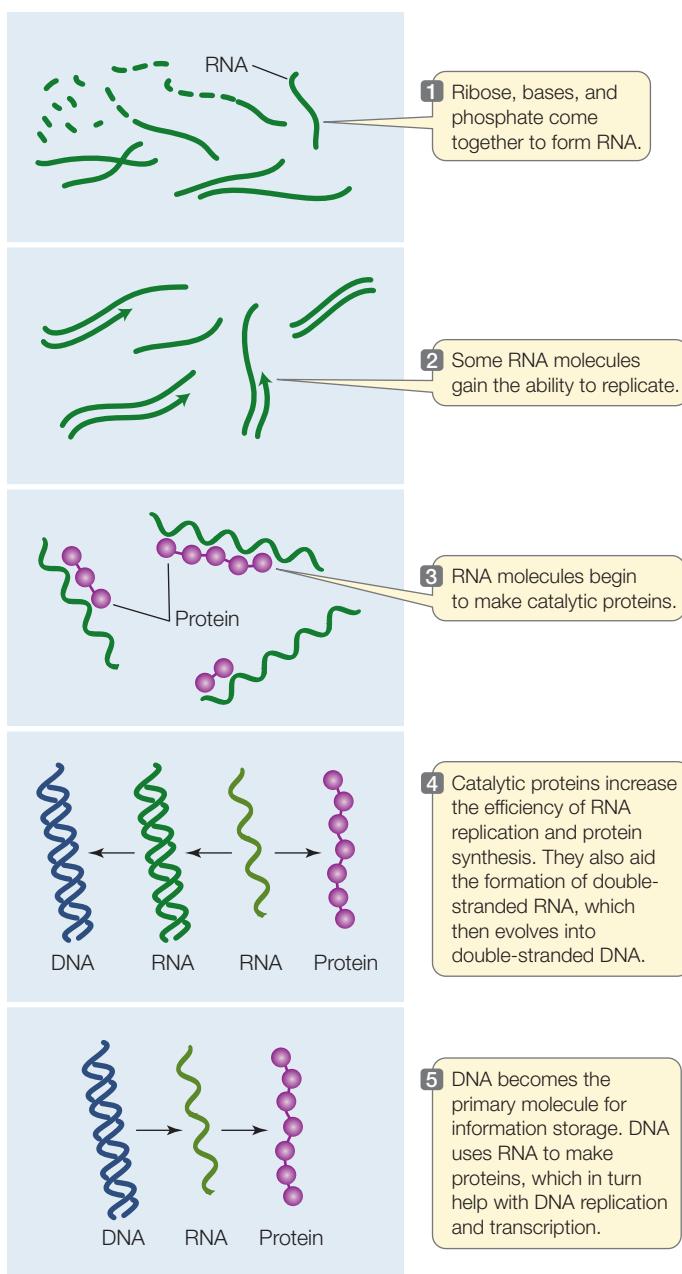
CHEMICAL CHANGES (METABOLISM) FIRST In this model, life began in tiny droplets, or compartments, that concentrated and sepa-



4.10 Two Pathways to Life Biologists have proposed two ways in which simple monomers could have become self-replicating systems capable of biological functions. (A) The chemical changes (metabolism) first pathway. (B) The replicator first pathway.

rated their contents from the external environment. Within such a chemically rich environment, some substances could occasionally and randomly undergo chemical changes. Proponents of this model speculate that those compartments where the changes were effective for survival in the environment might even have been selected for growth and some primitive form of reproduction. Could catalysis, the speeding up of reactions essential for life, occur in such an environment? The German scientist Günter Wächtershäuser proposed that catalysis and reproduction could have occurred without proteins on a mineral called pyrite (iron disulfide), which has been found at hydrothermal vents and which could serve as a source of energy for polymerization reactions. Over time, nucleic acids and eventually proteins might have formed in the concentrated droplets. Then, in some of these proteins, the ability to catalyze biochemical reactions—including the replication of nucleic acids—could have evolved.

REPLICATOR FIRST In this model, the genetic material—nucleic acids—came first. The nucleotide building blocks made by prebiotic chemistry came together to form polymers. Some of these polymers might have had the right shape to be catalytic so that they could reproduce themselves and catalyze other chemical transformations. Such transformations might have included the synthesis of proteins, just as RNA is translated into proteins in living organisms today (see Figure 4.5). Along the way, those molecules that were best adapted to the environment would survive and reproduce. Eventually they would have become incorporated into living cells.



4.11 The “RNA World” Hypothesis In a world before DNA, this view postulates that RNA alone was both the blueprint for protein synthesis and a catalyst for its own replication. Eventually, the more compact information storage molecules of DNA could have evolved from RNA.

4.12 An Early Catalyst for Life? In the laboratory, a ribozyme (a folded RNA molecule) can catalyze the polymerization of several short RNA strands into a longer molecule. Such a process could be a precursor for the copying of nucleic acids, which is essential for their replication and for gene expression.

- There are two major problems with the replicator first model:
- Nucleic acid polymers have not been observed in prebiotic chemistry simulations.
 - DNA, the genetic material in almost all current organisms, is not self-catalytic.

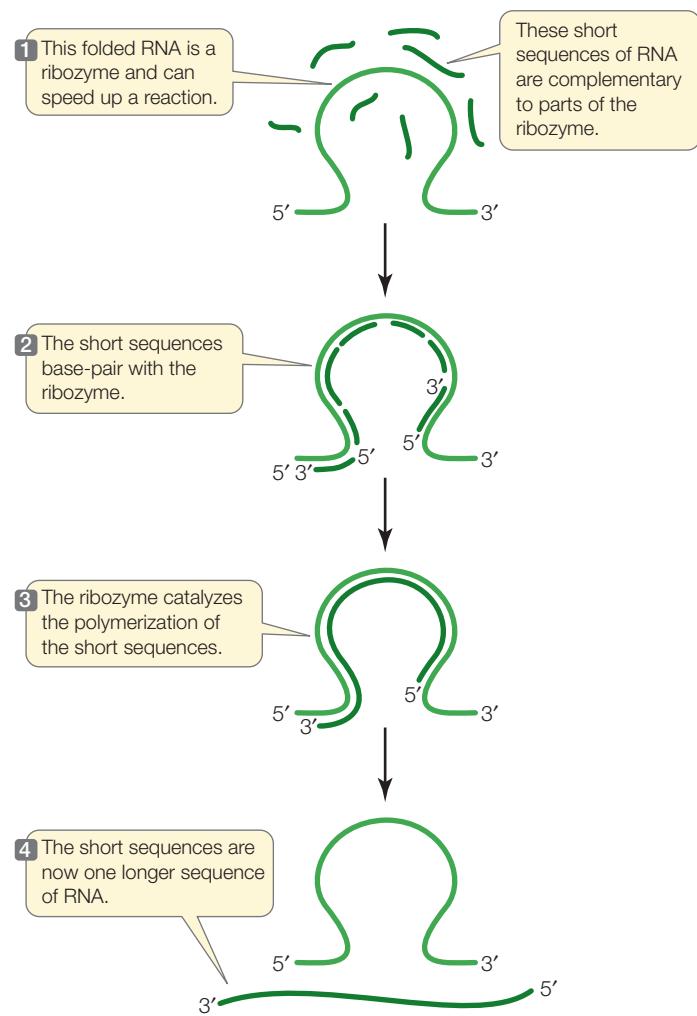
The first problem remains, but the second has a plausible solution: RNA can be a catalyst and can catalyze its own synthesis.

RNA may have been the first biological catalyst

The three-dimensional structure of a folded RNA molecule presents a unique surface to the external environment (see Figure 4.3). The surfaces of RNA molecules can be every bit as specific as those of proteins. Just as the shapes of proteins allow them to function as catalysts, speeding up reactions that would ordinarily take place too slowly to be biologically useful, the three-dimensional shapes and other chemical properties of certain RNA molecules allow them to function as catalysts. Catalytic RNAs, called **ribozymes**, can catalyze reactions on their own nucleotides as well as in other cellular substances. Although in retrospect it is not too surprising, the discovery of catalytic RNAs was a major shock to a community of biologists who were convinced that all biological catalysts were proteins (enzymes). It took almost a decade for the work of the scientists involved, Thomas Cech and Sidney Altman, to be fully accepted by other scientists. Later, they were awarded the Nobel Prize.

Given that RNA can be both informational (in its nucleotide sequence) and catalytic (due to its ability to form unique three-dimensional shapes), it has been hypothesized that early life consisted of an “RNA world”—a world before DNA. It is thought that when RNA was first made, it could have acted as a catalyst for its own replication as well as for the synthesis of proteins. DNA could eventually have evolved from RNA (Figure 4.11). Some laboratory evidence supports this scenario:

- When certain short RNA sequences are added to a mixture of nucleotides, RNA polymers can be formed at a rate 7 million times greater than the formation of polymers without the added RNA. This added RNA is not a template, but a catalyst.
- In the test tube, a ribozyme can catalyze the assembly of short RNAs into a longer molecule (Figure 4.12). This may be how nucleic acid replication evolved.
- In living organisms today, the formation of peptide linkages (see Figure 3.6) is catalyzed by ribozymes.
- In certain viruses called retroviruses, there is an enzyme called reverse transcriptase that catalyzes the synthesis of DNA from RNA.



4.3 RECAP

The emergence of the chemical reactions characteristic of life (metabolism), and the polymerization of monomers to polymers, may have occurred on the surfaces of hydrothermal vents. One theory proposes that metabolism came before polymerization; another suggests that the reverse occurred. RNA may have been the first genetic material and catalyst.

- What are the two theories for the emergence of metabolism and polymers? See pp. 69–71 and Figure 4.10
- How does RNA self-replicate? See p. 71 and Figure 4.12

The discovery of mechanisms for the formation of small and large molecules is essential to answering questions about the origin of life on Earth. But we also need to understand how organized systems formed that include these molecules and display the characteristic properties of life, such as reproduction, energy processing, and responsiveness to the environment. These properties are present in cells, and we now turn to ideas on their origin.

4.4 How Did the First Cells Originate?

As you have seen from many of the theories for the origin of life, the evolution of biochemistry occurred under localized conditions. That is, the chemical reactions of metabolism, polymerization, and replication could not occur in a dilute aqueous environment. There had to be a compartment of some sort that brought together and concentrated the compounds involved in these events. Biologists have proposed that initially this compartment may have simply been a tiny droplet of water on the surface of a rock. But another major event in the origin of life was necessary.

Life as we know it is separated from the environment within structurally defined units called **cells**. The internal contents of a cell are separated from the nonbiological environment by a special barrier—a **membrane**. The membrane is not just a barrier; it regulates what goes into and out of the cell, as we describe in Chapter 6. This role of the surface membrane is very important because it permits the interior of the cell to maintain a chemical composition that is different from its external environment. How did the first cells with membranes come into existence?

Experiments describe the origin of cells

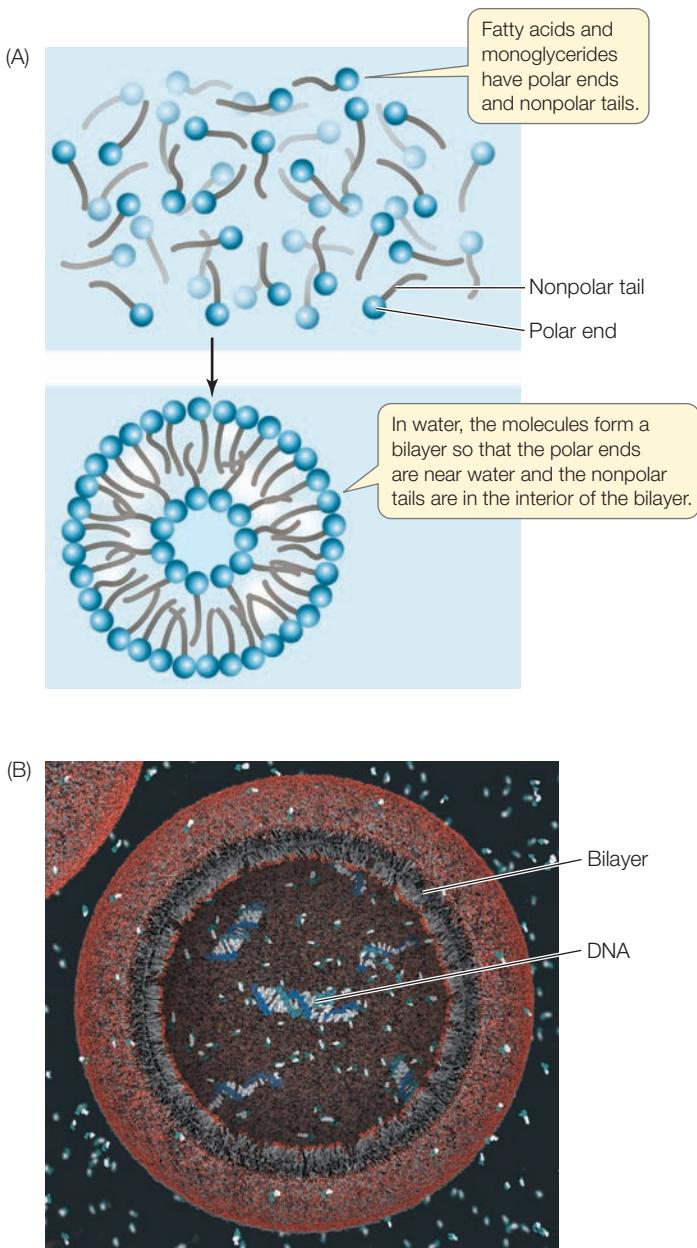
Jack Szostak and his colleagues at Harvard University built a laboratory model that gives insights into the origin of cells. To do this, they first put fatty acids (which can be made in prebiotic experiments) into water. Recall from Chapter 3 that fatty acids are *amphipathic*: they have a hydrophilic polar end and a long, nonpolar tail that is hydrophobic (see Figure 3.20). When placed in water, fatty acids will arrange themselves in a round “huddle” much like a football team: the hydrophilic ends point outward to interact with the aqueous environment and the fatty acid tails point inward, away from the water molecules.

What if some water becomes trapped in the interior of this “huddle”? Now the layer of hydrophobic fatty acid tails is in water, which is an unstable situation. To stabilize this, a second layer of fatty acids forms. This *lipid bilayer* has the polar ends of the fatty acids facing both outward and inward, because they are attracted to the polar water molecules present on each side of the double layer. The nonpolar tails form the interior of the bilayer (Figure 4.13). These prebiotic, water-filled structures, defined by a lipid bilayer membrane, very much resemble living cells. Scientists refer to these compartments as **protocells**. Examining their properties revealed that

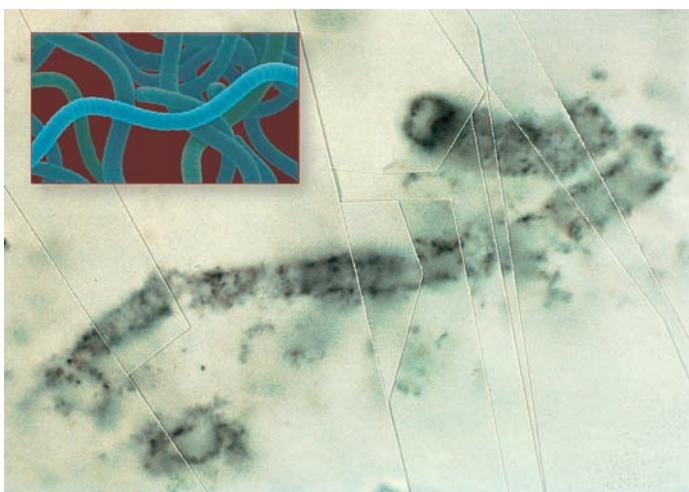
- Large molecules such as DNA or RNA could not pass through the bilayer to enter the protocells, but small molecules such as sugars and individual nucleotides could.
- Nucleic acids inside the protocells could replicate using the nucleotides from outside. When the investigators placed a short nucleic acid strand capable of self-replication inside protocells and added nucleotides to the watery environment outside, the nucleotides crossed the barrier, entered the protocells, and became incorporated into new polynucleotide

chains. This may have been the first step toward cell reproduction, and it took place without protein catalysis.

Were these protocells truly cells, and was the lipid bilayer produced in these experiments a true cell membrane? Certainly not. The protocells could not fully reproduce, nor could they carry out all the metabolic reactions that take place in modern cells. The simple lipid bilayer had few of the sophisticated functions



4.13 Protocells (A) In a series of experiments in the Szostak lab, researchers mixed fatty acid molecules in water. The molecules formed bilayers that have some of the properties of a cell membrane. The bilayers and the water “trapped” inside them are essential to form a protocell. (B) A model of the protocell. Nutrients and nucleotides (blue and white particles) pass through the “membrane” and enter the protocell, where they copy an already present DNA template. The new copies of DNA remain in the protocell.



4.14 The Earliest Cells? This fossil from Western Australia is 3.5 billion years old. Its form is similar to that of modern filamentous cyanobacteria (inset).

of modern cell membranes. Nevertheless, the protocell may be a reasonable facsimile of a cell as it evolved billions of years ago:

- It can act as a system of interacting parts
- It is capable of organization and self-catalysis
- It includes an interior that is distinct from the exterior environment.

These are all fundamental characteristics of living cells.

Some ancient cells left a fossil imprint

In the 1990s, scientists made an extremely rare find: a formation of ancient rocks in Australia that had remained relatively unchanged since they first formed 3.5 billion years ago. In one of these rock samples, geologist J. William Schopf of the University of California, Los Angeles, saw chains and clumps of what looked tantalizingly like contemporary cyanobacteria, or “blue-green” bacteria (**Figure 4.14**). Cyanobacteria are believed to

have been among the first organisms, because they can perform photosynthesis, converting CO_2 from the atmosphere and water into carbohydrates. Schopf needed to prove that the chains were once alive, not just the results of simple chemical reactions. He and his colleagues looked for chemical evidence of photosynthesis in the rock samples.

The use of carbon dioxide in photosynthesis is a hallmark of life and leaves a unique chemical signature—a specific ratio of isotopes of carbon ($^{13}\text{C}:\text{C}^{12}$) in the resulting carbohydrates. Schopf showed that the Australian material had this isotope signature. Furthermore, microscopic examination of the chains revealed *internal* substructures that are characteristic of living systems and were not likely to be the result of simple chemical reactions. Schopf’s evidence suggests that the Australian sample is indeed the remains of a truly ancient living organism.

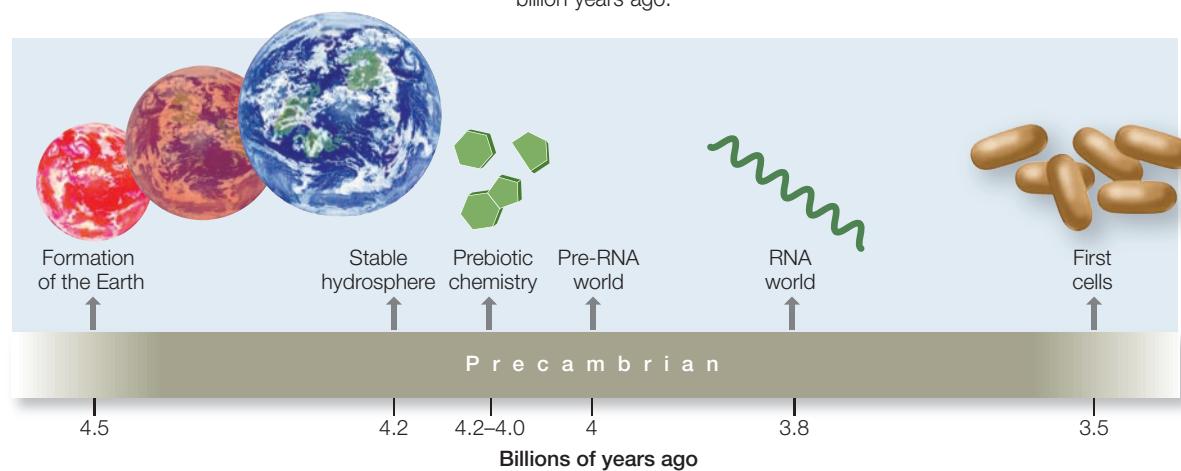
Taking geological, chemical, and biological evidence into account, it is plausible that it took about 500 million to a billion years from the formation of the Earth until the appearance of the first cells (**Figure 4.15**). Life has been cellular ever since. In the next chapter, we begin our study of cell structure and function.

4.4 RECAP

The chemical reactions that preceded living organisms probably occurred in specialized compartments, such as water droplets on the surfaces of minerals. Life as we know it did not begin until the emergence of cells. Protocells made in the laboratory have some of the properties of modern cells. Cell-like structures fossilized in ancient rocks date the first cells to about 3.5 billion years ago.

- Explain the importance of the cell membrane to the evolution of living organisms. **See p. 72**
- What is the evidence that ancient rocks contain the fossils of cells? **See p. 73**

4.15 The Origin of Life This highly simplified timeline gives a sense of the major events that culminated in the origin of life more than 3.5 billion years ago.



CHAPTER SUMMARY

4.1 What Are the Chemical Structures and Functions of Nucleic Acids?

- The unique function of the nucleic acids—DNA and RNA—is information storage. They form the hereditary material that passes genetic information from one generation to the next.
- Nucleic acids are polymers of nucleotides. A **nucleotide** consists of a phosphate group, a pentose sugar (**ribose** in RNA and **deoxyribose** in DNA), and a nitrogen-containing **base**. **Review Figure 4.1**
- In DNA, the nucleotide bases are **adenine**, **guanine**, **cytosine**, and **thymine**. **Uracil** replaces thymine in RNA. The nucleotides are joined by **phosphodiester linkages** between the sugar of one and the phosphate of the next, forming a nucleic acid polymer. **WEB ACTIVITY 4.1**
- DNA is a **double helix** with two separate strands in which there is **complementary base pairing** based on hydrogen bonds between adenine and thymine (A-T) and between guanine and cytosine (G-C). The two strands of the DNA double helix run in opposite directions. RNA consists of one chain of nucleotides. Hydrogen bonding can occur within the single strand of RNA, forming double-stranded regions and giving the molecule a three-dimensional surface shape. **Review Figures 4.2 and 4.3; WEB ACTIVITY 4.2**
- The information content of DNA and RNA resides in their **base sequences**.
- DNA is expressed as RNA in **transcription**. RNA can then specify the amino acid sequence of a protein in **translation**. **Review Figures 4.5 and 4.6**

4.2 How and Where did the Small Molecules of Life Originate?

- Historically, many cultures believed that life originates repeatedly by **spontaneous generation**. This was disproven experimentally. **Review Figure 4.7; ANIMATED TUTORIAL 4.1**
- Life probably originated from chemical reactions. A prerequisite for life is the presence of water.

- The presence of chemical traces of life on meteorites that have landed on Earth suggests that life might have originated extraterrestrially.
- Chemical experiments modeling the prebiotic conditions on Earth have shown that the small molecules that characterize life could have been formed from atmospheric chemicals. **Review Figure 4.9; ANIMATED TUTORIAL 4.2**

4.3 How Did the Large Molecules of Life Originate?

- Polymerization of small molecules to polymers could occur in small compartments such as droplets or on surfaces. Both of these conditions concentrate molecules such that reactions are favored.
- The “metabolism first” theory of polymerization proposes that chemical reactions involving small molecules evolved first, and some of them formed polymers that acted as genetic information and catalysts.
- The “replicator first” theory proposes that RNA formed early, and acted as both genetic material and catalyst. Then reactions involving small molecules could occur. **Review Figure 4.10**
- In contemporary organisms, RNA can act as both an information molecule and as a catalyst. This favors the replicator first model. The **RNA world** may have been an important step on the way to life. **Review Figure 4.11**

4.4 How Did Cells Originate?

- A key to the emergence of living cells was the prebiotic chemical generation of compartments enclosed by **membranes**. Such enclosed compartments permitted the generation and maintenance of internal chemical conditions that were different from those in the exterior environment.
- In the laboratory, fatty acids and related lipids assemble into **protocells** that have some of the characteristics of cells. **Review Figure 4.13**
- Ancient rocks (3.5 billion years old) have been found with imprints that are probably fossils of early cells.

SELF-QUIZ

- A nucleotide in DNA is made up of
 - four bases.
 - a base plus a ribose sugar.
 - a base plus a deoxyribose sugar plus phosphate.
 - a sugar plus a phosphate.
 - a sugar and a base.
- Nucleotides in RNA are connected to one another in the polynucleotide chain by
 - covalent bonds between bases.
 - covalent bonds between sugars.
 - covalent bonds between sugar and phosphate.
 - hydrogen bonds between purines.
 - hydrogen bonds between any bases.
- Which is a difference between DNA and RNA?
 - DNA is single-stranded and RNA is double-stranded.
 - DNA is only informational and RNA is only catalytic.
 - DNA contains deoxyribose and RNA contains ribose.
 - DNA is transcribed and RNA is replicated.
 - DNA contains uracil (U) and RNA contains thymine (T).
- The nucleotide sequence of DNA
 - is the same in all organisms of a species.
 - contains only information for translation.
 - evolved before RNA.
 - contains the four bases, A, T, G, and C.
 - is produced by prebiotic chemistry experiments.
- Spontaneous generation of life from nonliving materials
 - can occur in dark places.
 - has not been a belief of humans.
 - has never occurred.
 - requires only nucleotides and fatty acids.
 - was disproven for microorganisms by Pasteur’s experiment.
- The components in the atmosphere for the Miller–Urey experiment on prebiotic synthesis did not include
 - H_2 .
 - H_2O .
 - O_2 .
 - NH_3 .
 - CH_4 .

7. All of the major building blocks of macromolecules were made in Miller–Urey prebiotic synthesis experiments *except*
 - a. amino acids.
 - b. hexose sugars.
 - c. bases for nucleotides.
 - d. fatty acids.
 - e. ribose.
8. The “RNA world” hypothesis proposes that
 - a. RNA formed from DNA.
 - b. RNA was both a catalyst and genetic material.
 - c. RNA was a catalyst only.
 - d. RNA formed after proteins.
 - e. DNA formed after RNA was broken down.
9. Ribozymes are
 - a. enzymes that are made up of ribose sugar.
 - b. ancient catalysts that no longer exist.
 - c. RNA catalysts.
 - d. present in bacterial cells only.
 - e. less active than protein enzymes.
10. Findings in ancient rocks indicate cells first appeared
 - a. about 4.5 billion years ago.
 - b. about 3.5 billion years ago.
 - c. about 2 billion years ago.
 - d. before rocks were formed.
 - e. before water arrived on Earth.

FOR DISCUSSION

1. Are the statements “all life comes from pre-existing life” and “life on Earth could have arisen from prebiotic molecules” truly paradoxical? What conditions existing on Earth today might preclude the origin of life from such molecules?
2. Why might RNA have preceded proteins in the evolution of biological macromolecules?
3. Do you consider the two alternative theories presented in this chapter as possible explanations of the origin of life on Earth (that life came from outside of Earth, or that life arose on Earth through chemical evolution) to be equally plausible? Which do you favor, and why?
4. Why was the evolution of a self-contained cell essential for life as we know it?

ADDITIONAL INVESTIGATION

1. The interpretation of Pasteur’s experiment (see Figure 4.7) depended on the inactivation of microorganisms by heat. We now know of microorganisms that can survive extremely high temperatures (see Chapter 26). Does this change the interpretation of Pasteur’s experiment? What experiments would you do to inactivate such microbes?
2. The Miller–Urey experiment (see Figure 4.9) showed that it was possible for amino acids to be formed from gases that were hypothesized to have been in Earth’s early atmosphere. These amino acids were dissolved in water. Knowing what you do about the polymerization of amino acids into proteins (see Figure 3.6), how would you set up experiments to show that proteins can form under the conditions of early Earth? What properties would you expect of those proteins?

WORKING WITH DATA (GO TO yourBioPortal.com)

Synthesis of Prebiotic Molecules in an Experimental Atmosphere

In this hands-on exercise, you will examine the original research paper of Miller and Urey to see the experimental approach they used to show that amino acids could be made in a simulation of Earth’s early atmosphere (Figure 4.9). You will also analyze more recent data using the same apparatus.

Disproving the Spontaneous Generation of Life

In this hands-on exercise, you will examine data from an experiment similar to Pasteur’s famous experiments (Figure 4.7). By calculating growth rates in the different flasks, you will be able to see how Pasteur came to the conclusion he did.

Cells: The Working Units of Life

How to mend a broken heart

It is a day in the not-too-distant future. Decades of eating fatty foods, combined with an inherited tendency to deposit cholesterol in his arteries, have finally caught up with 70-year-old Don. A blood clot has closed off blood flow to part of his heart, leading to a heart attack and severe damage to that vital organ.

If this had happened today, Don would have been faced with a long period of rehabilitation, taking medications to manage his weakened heart. Instead, his physicians take a pinch of skin tissue from his arm and bring it to a laboratory. After certain DNA sequences are added, Don's skin cells no longer look and act like skin cells: They are undifferentiated (unspecialized) and reproduce continuously in the laboratory dish. These cells are also multipotent stem

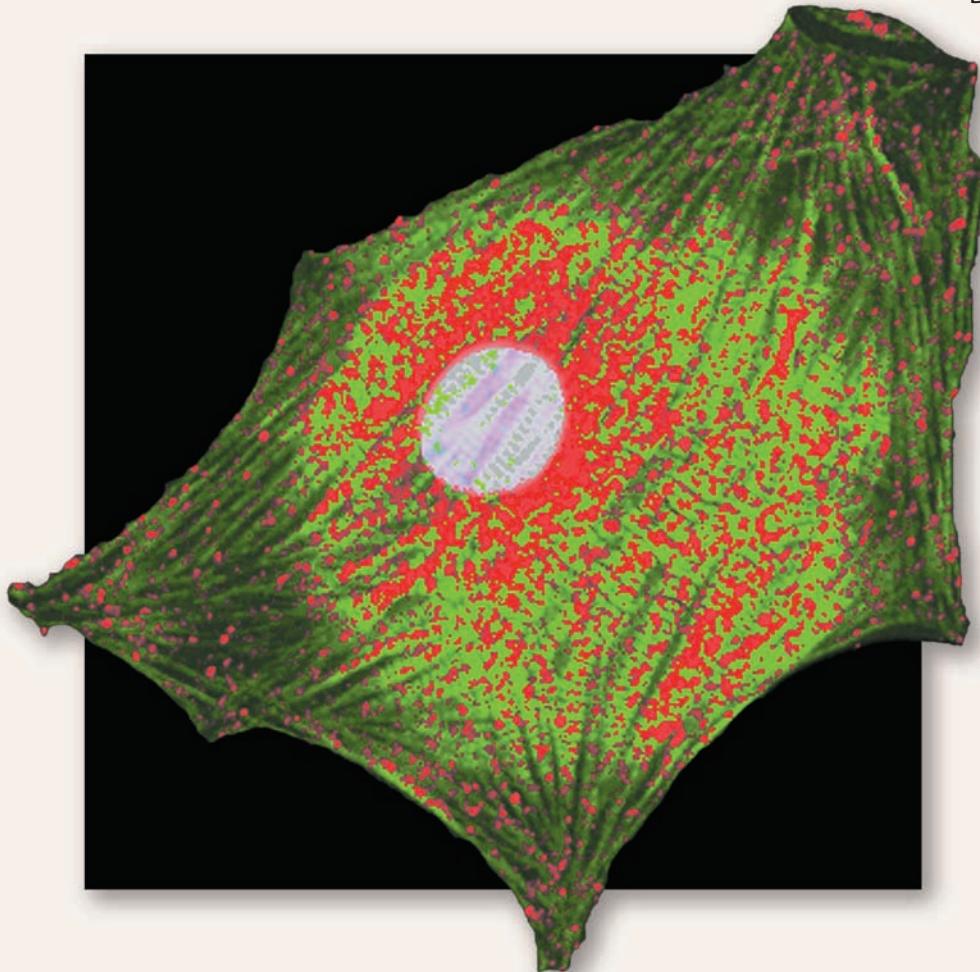
cells, able to differentiate into almost any type of cell in the body if given the right environment. When they are injected directly into Don's heart, his stem cells soon become heart muscle cells, repairing the damage caused by the heart attack. Don leaves the hospital with full cardiac function and recommendations for a healthy diet.

You are probably familiar with another type of multipotent cell, the fertilized human egg. This single cell ultimately produces the tens of trillions of cells that make up the human body. The fertilized egg is programmed to generate an entire organism—not just the heart and skin, but blood, nerves, liver, brain, and even bones—and for this reason is called totipotent (“toti” means all; “multi” means most). In contrast, the stem cells derived from

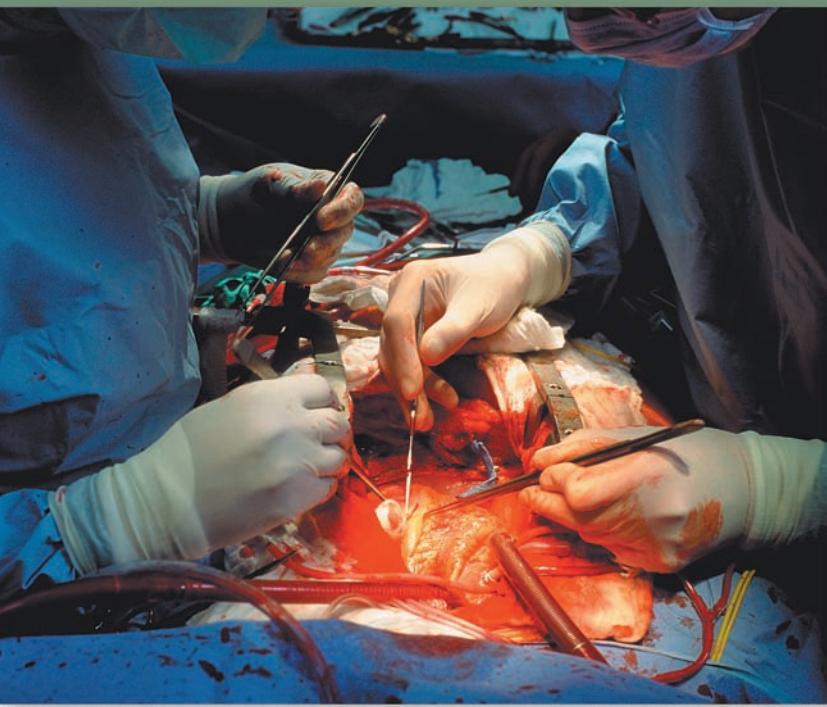
Don's skin need specific external signals to differentiate into other kinds of cells, and could not develop into an entire person.

The potential uses of stem cells in medicine have generated a lot of excitement in recent years. Such widely read periodicals as *Time* have hailed advances in stem cell research as “breakthroughs of the year.” Patients with the neurological disorder Parkinson’s disease dream of the day when their skin cells can be turned into brain cells to fix their damaged nervous systems. People with diabetes hope for stem cells to repair their pancreases. The list is long.

Behind all of this hope and the research it inspires is a cornerstone of biological science: the cell theory. As you saw in the last



A New Heart Cell This cardiac stem cell is developing into a fully differentiated heart cell. The hope is to be able to coax stem cells to follow this path or to produce other cell types to repair damaged tissues.



Open Heart Surgery Stem cell therapies may provide alternative approaches to treating heart disease in the future.

chapter, a key event in the emergence of life was the enclosure of biochemical reactions inside a cell, thus concentrating them and separating them from the external environment. These are the first two tenets of the cell theory, that the cell is the unit of life and that the activities of life either happen inside cells or are caused by them. Don's stem cells contain not just the activities of a living entity, but also the potential to change those activities in new directions. The third tenet of the cell theory—equally important—is that the cell is the unit of reproduction: all cells come from pre-existing cells. Stem cell therapy does not create new cells out of thin air; it coaxes existing ones to differentiate and reproduce along the desired path.

IN THIS CHAPTER we examine the structure and some of the functions of cells. We will begin with a fuller explanation of cell theory. Then, we will examine the relatively simple cells of prokaryotes. This is followed by a tour of the more complex eukaryotic cell and its various internal compartments, each of which performs specific functions. Finally, we discuss ideas on how complex cells evolved.

CHAPTER OUTLINE

- 5.1 What Features Make Cells the Fundamental Units of Life?
- 5.2 What Features Characterize Prokaryotic Cells?
- 5.3 What Features Characterize Eukaryotic Cells?
- 5.4 What Are the Roles of Extracellular Structures?
- 5.5 How Did Eukaryotic Cells Originate?

5.1 What Features Make Cells the Fundamental Units of Life?

In Chapter 1 we introduced some of the characteristics of life: chemical complexity, growth and reproduction, the ability to refashion substances from the environment, and the ability to determine what substances can move into and out of the organism. These characteristics are all demonstrated by cells. Just as atoms are the building blocks of chemistry, cells are the building blocks of life.

The **cell theory** is described in Section 1.1 as the first unifying principle of biology. There are three critical components of the cell theory:

- Cells are the fundamental units of life.
- All living organisms are composed of cells.
- All cells come from preexisting cells.

Cells contain water and the other small and large molecules, which we examined in Chapters 2–4. Each cell contains at least 10,000 different types of molecules, most of them present in many copies. Cells use these molecules to transform matter and energy, to respond to their environments, and to reproduce themselves.

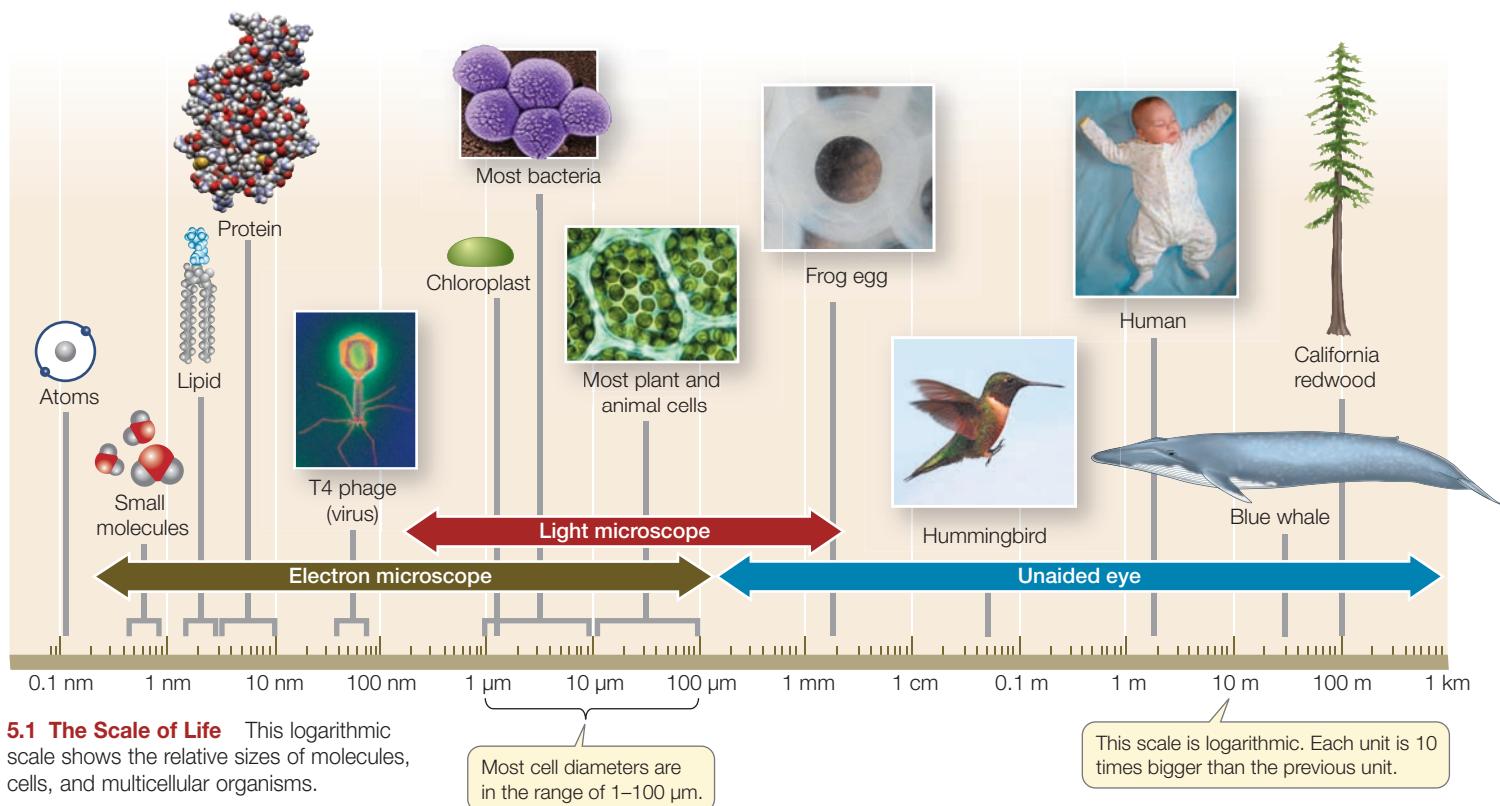
The cell theory has three important implications:

- Studying cell biology is in some sense the same as studying life. The principles that underlie the functions of the single cell of a bacterium are similar to those governing the approximately 60 trillion cells of your body.
- Life is continuous. All those cells in your body came from a single cell, a fertilized egg. That egg came from the fusion of two cells, a sperm and an egg, from your parents. The cells of your parents' bodies were all derived from their parents, and so on back through generations and evolution to the first living cell.
- The origin of life on Earth was marked by the origin of the first cells (see Chapter 4).

Even the largest creatures on Earth are composed of cells, but the cells themselves are usually too small for the naked eye to see. Why are cells so small?

Cell size is limited by the surface area-to-volume ratio

Most cells are tiny. In 1665, the early microscopist Robert Hooke estimated that in one square inch of cork, which he examined under his magnifying lens, there were 1,259,712,000 cells! The volumes of cells range from 1 to 1,000 cubic micrometers (μm^3). There are some exceptions: the eggs of birds are single cells that



yourBioPortal.com

GO TO Web Activity 5.1 • The Scale of Life

are, relatively speaking, enormous, and individual cells of several types of algae and bacteria are large enough to be viewed with the unaided eye (**Figure 5.1**). And although neurons (nerve cells) have volumes that are within the “usual” range, they often have fine projections that may extend for meters, carrying signals from one part of a large animal to another. So there is enormous diversity among cells in their dimensions and volumes, but cells are usually very small.

Small cell size is a practical necessity arising from the change in the **surface area-to-volume ratio** of any object as it increases in size. As an object increases in volume, its surface area also increases, but not at the same rate (**Figure 5.2**). This phenomenon has great biological significance for two reasons:

- The volume of a cell determines the amount of chemical activity it carries out per unit of time.
- The surface area of a cell determines the amount of substances that can enter it from the outside environment, and the amount of waste products that can exit to the environment.

5.2 Why Cells Are Small Whether it is cuboid (A) or spheroid (B), as an object grows larger its volume increases more rapidly than its surface area. Cells must maintain a large surface area-to-volume ratio in order to function. This fact explains why large organisms must be composed of many small cells rather than a few huge ones.

(A) Cubes

			Smaller surface area compared to volume.
			Larger surface area compared to volume.
1-mm cube	2-mm cube	4-mm cube	
Surface area $6 \text{ sides} \times 1^2 = 6 \text{ mm}^2$	6 sides $\times 2^2 = 24 \text{ mm}^2$	6 sides $\times 4^2 = 96 \text{ mm}^2$	
Volume $1^3 = 1 \text{ mm}^3$	$2^3 = 8 \text{ mm}^3$	$4^3 = 64 \text{ mm}^3$	
Surface area-to-volume ratio 6:1	3:1	1.5:1	

(B) Spheres

Diameter 1 μm	2 μm	3 μm	
Surface area $4\pi r^2$	$3.14 \mu\text{m}^2$	$12.56 \mu\text{m}^2$	$28.26 \mu\text{m}^2$
Volume $\frac{4}{3}\pi r^3$	$0.52 \mu\text{m}^3$	$4.19 \mu\text{m}^3$	$14.18 \mu\text{m}^3$
Surface area-to-volume ratio 6:1	3:1	2:1	

As a living cell grows larger, its chemical activity, and thus its need for resources and its rate of waste production, increases faster than its surface area. (The surface area increases in proportion to the square of the radius, while the volume increases much more—in proportion to the cube of the radius.) In addition, substances must move from one site to another within the cell; the smaller the cell, the more easily this is accomplished. This explains why large organisms must consist of many small cells: cells must be small in volume in order to maintain a large enough surface area-to-volume ratio and an ideal internal volume. The large surface area represented by the many small cells of a multicellular organism enables it to carry out the many different functions required for survival.

Microscopes reveal the features of cells

Microscopes do two different things to allow cells and details within them to be seen by the human eye. First, they increase the apparent size of the object: this is called *magnification*. But just increasing the magnification does not necessarily mean that the object will be seen clearly. In addition to being larger, a magnified object must be sharp, or clear. This is a property called *resolution*. Formally defined, resolution is the minimum distance two objects can be apart and still be seen as two objects. Resolution for the human eye is about 0.2 mm (200 μm). Most cells are much smaller than 200 μm , and thus are invisible to the human eye. Microscopes magnify and increase resolution so that cells and their internal structures can be seen clearly (**Figure 5.3**).

There are two basic types of microscopes—*light microscopes* and *electron microscopes*—that use different forms of radiation (see Figure 5.3). While the resolution is better in electron microscopy, we should emphasize that because cells are prepared in a vacuum, only dead, dehydrated cells are visualized. Therefore, the preparation of cells for electron microscopy may alter them, and this must be taken into consideration when interpreting the images produced. On the other hand, light microscopes can be used to visualize living cells (for example, by phase-contrast microscopy; see Figure 5.3).

Before we delve into the details of cell structure, it is useful to consider the many uses of microscopy. An entire branch of medicine, *pathology*, makes use of many different methods of microscopy to aid in the analysis of cells and the diagnosis of diseases. For instance, a surgeon might remove from a body some tissue suspected of being cancerous. The pathologist might:

- examine the tissue quickly by phase-contrast microscopy or interference-contrast microscopy to determine the size, shape, and spread of the cells
- stain the tissue with a general dye and examine it by bright-field microscopy to bring out features such as the shape of the nucleus, or cell division characteristics
- stain the tissue with a fluorescent dye and examine it by fluorescence microscopy or confocal microscopy for the presence of specific proteins that are diagnostic of a particular cancer
- examine the tissue under the electron microscope to observe its most minute internal structures, such as the shapes

of the mitochondria and the chromatin. (These structures are described in Section 5.3.)

The plasma membrane forms the outer surface of every cell

While the structural diversity of cells can often be observed using light microscopy, the **plasma membrane** is best observed with an electron microscope. This very thin structure forms the outer surface of every cell, and it has more or less the same thickness and molecular structure in all cells. Biochemical methods have shown that membranes have great functional diversity. These methods have revealed that the thin, almost invisible plasma membrane is actively involved in many cellular functions—it is not a static structure. The plasma membrane separates the interior of the cell from its outside environment, creating a segregated (but not isolated) compartment. The presence of this outer limiting membrane is a feature of all cells. What is the composition and molecular architecture of this amazing structure?

The plasma membrane is composed of a *phospholipid bilayer* (or simply *lipid bilayer*), with the hydrophilic “heads” of the lipids facing the cell’s aqueous interior on one side of the membrane and the extracellular environment on the other (see Figure 3.20). Proteins and other molecules are embedded in the lipids. The membrane is not a rigid, static structure. Rather, it is an oily fluid, in which the proteins and lipids are in constant motion. This allows the membrane to move and change the shape of the cell. A detailed description of the structure and functions of the plasma membrane is given in Chapter 6. Here is a brief summary:

- The plasma membrane acts as a *selectively permeable barrier*, preventing some substances from crossing it while permitting other substances to enter and leave the cell. For example, macromolecules such as DNA and proteins cannot normally cross the plasma membrane, but some smaller molecules such as oxygen can. In addition to size, other factors such as polarity determine a molecule’s ability to cross the plasma membrane: because the membrane is composed mostly of hydrophobic fatty acids, nonpolar molecules cross it more easily than polar or charged molecules.
- The plasma membrane allows the cell to maintain a more or less *constant internal environment*. A self-maintaining, constant internal environment (known as *homeostasis*) is a key characteristic of life that will be discussed in detail in Chapter 40. One way that the membrane does this is by actively regulating the transport of substances across it. This dynamic process is distinct from the more passive process of diffusion, which is dependent on the size of a molecule.
- As the cell’s boundary with the outside environment, the plasma membrane is important in *communicating* with adjacent cells and receiving signals from the environment. We will describe this function in Chapter 7.
- The plasma membrane often has proteins protruding from it that are responsible for *binding* and *adhering* to adjacent

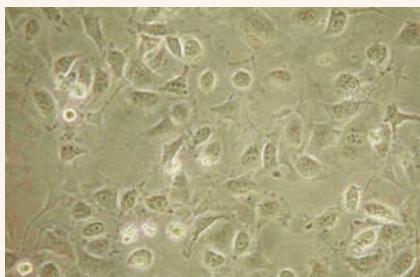
TOOLS FOR INVESTIGATING LIFE

5.3 Looking at Cells

The six images on this page show some techniques used in light microscopy. The three images on the following page were created using electron microscopes. All of these images are of a particular type of cultured cell known as HeLa cells. Note that the images in most cases are flat, two-dimensional views. As you look at images of cells, keep in mind that they are three-dimensional structures.



In a **light microscope**, glass lenses and visible light are used to form an image. The resolution is about $0.2 \mu\text{m}$, which is 1,000 times greater than that of the human eye. Light microscopy allows visualization of cell sizes and shapes and some internal cell structures. Internal structures are hard to see under visible light, so cells are often chemically treated and stained with various dyes to make certain structures stand out by increasing contrast.



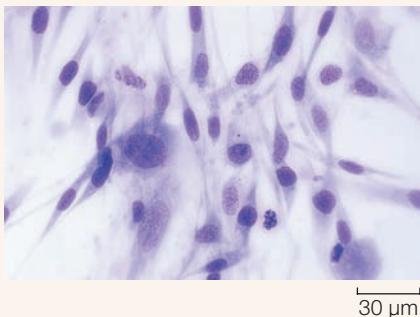
In **bright-field microscopy**, light passes directly through these human cells. Unless natural pigments are present, there is little contrast and details are not distinguished.



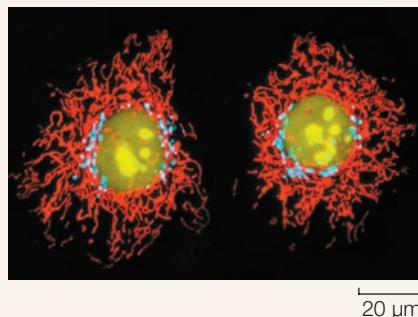
In **phase-contrast microscopy**, contrast in the image is increased by emphasizing differences in refractive index (the capacity to bend light), thereby enhancing light and dark regions in the cell.



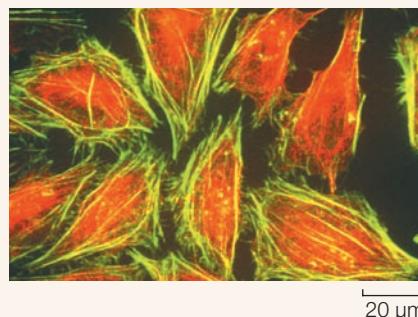
Differential interference-contrast microscopy uses two beams of polarized light. The combined images look as if the cell is casting a shadow on one side.



In **stained bright-field microscopy**, a stain enhances contrast and reveals details not otherwise visible. Stains differ greatly in their chemistry and their capacity to bind to cell materials, so many choices are available.



In **fluorescence microscopy**, a natural substance in the cell or a fluorescent dye that binds to a specific cell material is stimulated by a beam of light, and the longer-wavelength fluorescent light is observed coming directly from the dye.



Confocal microscopy uses fluorescent materials but adds a system of focusing both the stimulating and emitted light so that a single plane through the cell is seen. The result is a sharper two-dimensional image than with standard fluorescence microscopy.

cells. Thus the plasma membrane plays an important structural role and contributes to cell shape.

All cells are classified as either prokaryotic or eukaryotic

As we learned in Section 1.2, biologists classify all living things into three domains: Archaea, Bacteria, and Eukarya. The organ-

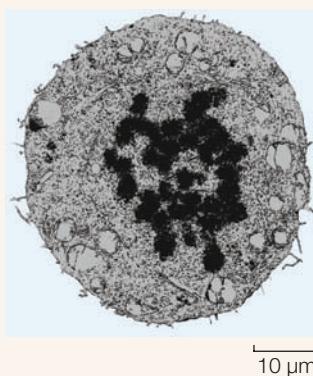
isms in Archaea and Bacteria are collectively called **prokaryotes** because they have in common a prokaryotic cell organization. A prokaryotic cell does not typically have membrane-enclosed internal compartments; in particular, it does not have a nucleus. The first cells were probably similar in organization to those of modern prokaryotes.

TOOLS FOR INVESTIGATING LIFE

5.3 Looking at Cells (continued)

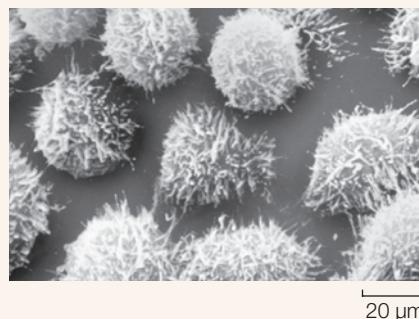


In an *electron microscope*, electromagnets are used to focus an electron beam, much as a light microscope uses glass lenses to focus a beam of light. Since we cannot see electrons, the electron microscope directs them through a vacuum at a fluorescent screen or photographic film to create a visible image. The resolution of electron microscopes is about 2 nm, which is about 100,000 times greater than that of the human eye. This resolution permits the details of many subcellular structures to be distinguished.



10 µm

In **transmission electron microscopy (TEM)**, a beam of electrons is focused on the object by magnets. Objects appear darker if they absorb the electrons. If the electrons pass through they are detected on a fluorescent screen.



20 µm

Scanning electron microscopy (SEM) directs electrons to the surface of the sample, where they cause other electrons to be emitted. These electrons are viewed on a screen. The three-dimensional surface of the object can be visualized.



0.1 µm

In **freeze-fracture microscopy**, cells are frozen and then a knife is used to crack them open. The crack often passes through the interior of plasma and internal membranes. The “bumps” that appear are usually large proteins or aggregates embedded in the interior of the membrane.

yourBioPortal.com

GO TO Web Activity 5.2 • Know Your Techniques

Eukaryotic cell organization, on the other hand, is found in members of the domain Eukarya (**eukaryotes**), which includes the protists, plants, fungi, and animals. As we will discuss later in this chapter, eukaryotic cells probably evolved from prokaryotes. In contrast to the prokaryotes, the genetic material (DNA) of eukaryotic cells is contained in a special membrane-enclosed compartment called the **nucleus**. Eukaryotic cells also contain other membrane-enclosed compartments in which specific chemical reactions occur. For example, some of the key reactions that generate usable chemical energy for cells take place in mitochondria. The internal membranes that enclose these compartments have the same basic composition, structure and properties as the plasma membrane. The efficiency afforded by these compartments has led to the impressive functions that can occur in eukaryotic cells, and their specialization into tissues as diverse as the parts of a flower, muscles, and nerves.

5.1 RECAP

The cell theory is a unifying principle of biology. Surface area-to-volume ratios limit the sizes of cells. Both prokaryotic and eukaryotic cells are enclosed within a plasma membrane, but prokaryotic cells lack the membrane-enclosed internal compartments found in eukaryotes.

- How does cell biology embody all the principles of life? **See p. 77**
- Why are cells small? **See pp. 77–79 and Figure 5.2**
- Explain the importance of the plasma membrane to cells. **See pp. 79–80**

As we mentioned in this section, there are two structural themes in cell architecture: prokaryotic and eukaryotic. We now turn to the organization of prokaryotic cells.

5.2 What Features Characterize Prokaryotic Cells?

Prokaryotes can derive energy from more diverse sources than any other living organisms. They can tolerate environmental extremes—such as very hot springs with temperatures up to 100°C (*Thermus aquaticus*) or very salty water (*Halobacterium*)—that would kill other organisms. As we examine prokaryotic cells in this section, bear in mind that there are vast numbers of prokaryotic species, and that the Bacteria and Archaea are distinguished in numerous ways. These differences, and the vast diversity of organisms in these two domains, will be the subject of Chapter 26.

The volume of a prokaryotic cell is generally about one fiftieth of the volume of a eukaryotic cell. Prokaryotic cells range from about 1 to 10 μm in length or diameter. Each individual prokaryote is a single cell, but many types of prokaryotes are usually seen in chains or small clusters, and some occur in large clusters containing hundreds of cells. In this section we will first consider the features shared by cells in the domains Bacteria and Archaea. Then we will examine structural features that are found in some, but not all, prokaryotes.

Prokaryotic cells share certain features

All prokaryotic cells have the same basic structure (Figure 5.4):

- The plasma membrane encloses the cell, regulating the traffic of materials into and out of the cell, and separating its interior from the external environment.

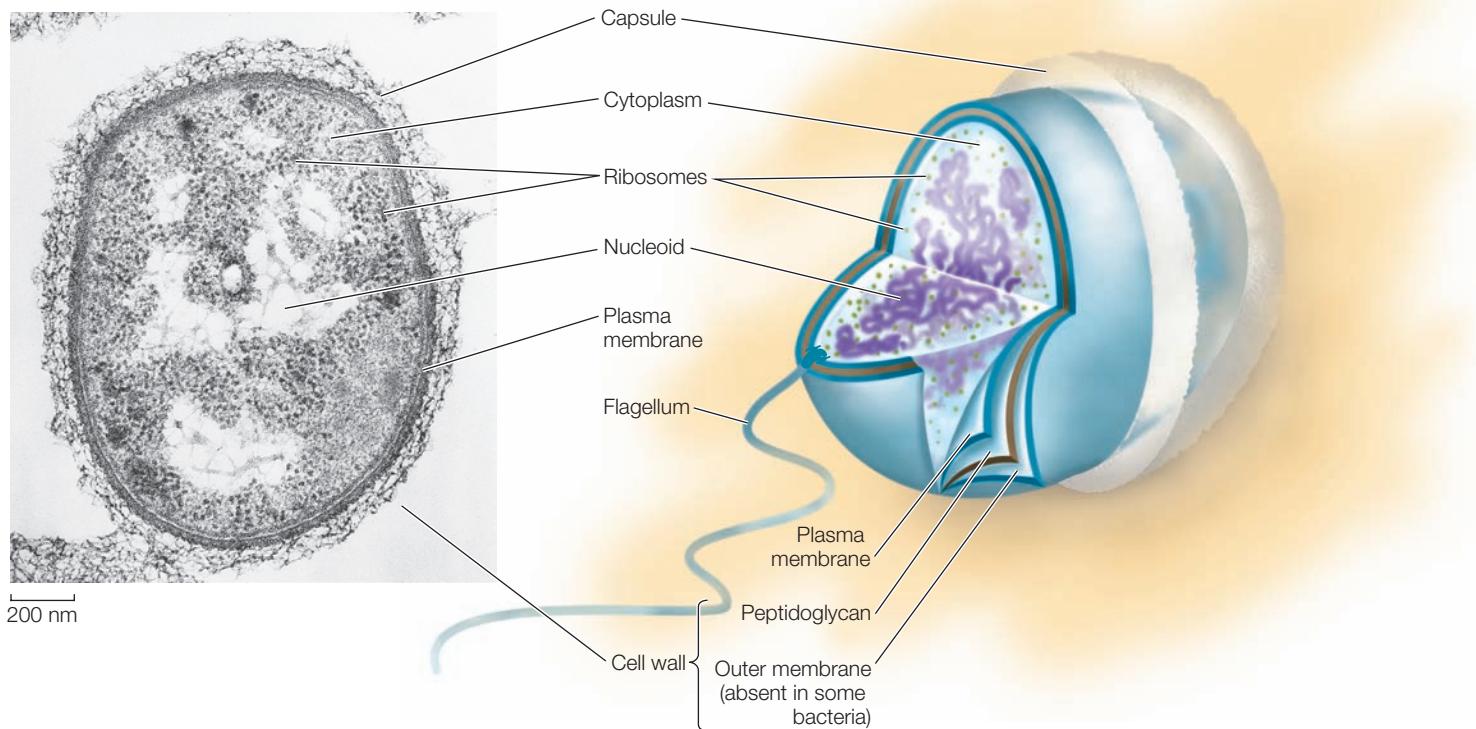
- The **nucleoid** is a region in the cell where the DNA is located. As we described in Section 4.1, DNA is the hereditary material that controls cell growth, maintenance, and reproduction.

The rest of the material enclosed in the plasma membrane is called the **cytoplasm**. The cytoplasm has two components: the cytosol and insoluble suspended particles, including ribosomes:

- The **cytosol** consists mostly of water that contains dissolved ions, small molecules, and soluble macromolecules such as proteins.
- **Ribosomes** are complexes of RNA and proteins that are about 25 nm in diameter. They can only be visualized with the electron microscope. They are the sites of protein synthesis, where information coded for in nucleic acids directs the sequential linking of amino acids to form proteins.

The cytoplasm is not a static region. Rather, the substances in this environment are in constant motion. For example, a typical protein moves around the entire cell within a minute, and it collides with many other molecules along the way.

Although they are structurally less complex than eukaryotic cells, prokaryotic cells are functionally complex, carrying out thousands of biochemical reactions. Based on our current knowledge about the origins of the first cells (see Section 4.4), some prokaryotic cell lineages must stretch back in time for more than 3 billion years. Thus, prokaryotes are very successful organisms from an evolutionary perspective.



5.4 A Prokaryotic Cell The bacterium *Pseudomonas aeruginosa* illustrates the typical structures shared by all prokaryotic cells. This bacterium also has a protective outer membrane that not all prokaryotes have. The flagellum and capsule are also structures found in some, but not all, prokaryotic cells.

Specialized features are found in some prokaryotes

As they evolved, some prokaryotes developed specialized structures that gave a selective advantage to those that had them: cells with these structures were better able to survive and reproduce in particular environments than cells lacking them. These structures include a protective cell wall, an internal membrane for compartmentalization of some chemical reactions, flagella for cell movement through the watery environment, and a rudimentary internal skeleton.

CELL WALLS Most prokaryotes have a cell wall located outside the plasma membrane. The rigidity of the cell wall supports the cell and determines its shape. The cell walls of most bacteria, but not archaea, contain peptidoglycan, a polymer of amino sugars that are cross-linked by covalent bonds to peptides, to form a single giant molecule around the entire cell. In some bacteria, another layer, the outer membrane (a polysaccharide-rich phospholipid membrane), encloses the peptidoglycan layer (see Figure 5.4). Unlike the plasma membrane, this outer membrane is not a major barrier to the movement of molecules across it.

Enclosing the cell wall in some bacteria is a slimy layer composed mostly of polysaccharides, and referred to as a capsule. In some cases these capsules protect the bacteria from attack by white blood cells in the animals they infect. Capsules also help to keep the cells from drying out, and sometimes they help bacteria attach to other cells. Many prokaryotes produce no capsule, and those that do have capsules can survive even if they lose them, so the capsule is not essential to prokaryotic life.

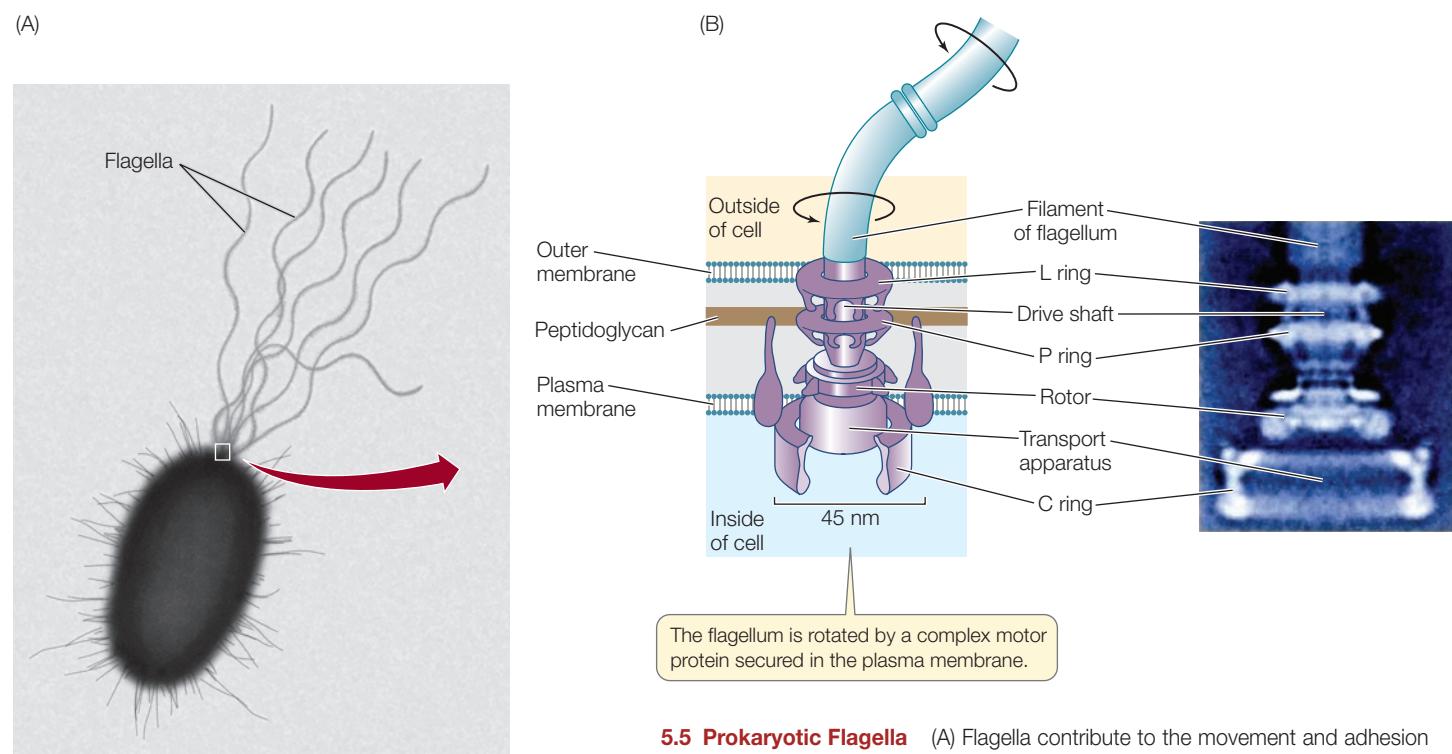
As you will see later in this chapter, eukaryotic plant cells also have a cell wall, but it differs in composition and structure from the cell walls of prokaryotes.

INTERNAL MEMBRANES Some groups of bacteria—including the cyanobacteria—carry out photosynthesis: they use energy from the sun to convert carbon dioxide and water into carbohydrates. These bacteria have an internal membrane system that contains molecules needed for photosynthesis. The development of photosynthesis, which requires membranes, was an important event in the early evolution of life on Earth. Other prokaryotes have internal membrane folds that are attached to the plasma membrane. These folds may function in cell division or in various energy-releasing reactions.

FLAGELLA AND PILI Some prokaryotes swim by using appendages called **flagella**, which sometimes look like tiny corkscrews (Figure 5.5A). In bacteria a single flagellum is made of a protein called flagellin. A complex motor protein spins the flagellum on its axis like a propeller, driving the cell along. The motor protein is anchored to the plasma membrane and, in some bacteria, to the outer membrane of the cell wall (Figure 5.5B).

We know that the flagella cause the motion of cells because if they are removed, the cells do not move.

Pili are structures made of protein that project from the surfaces of some types of bacterial cells. These hairlike structures are shorter than flagella, and are used for adherence. The sex-pili help bacteria join to one another to exchange genetic material. The *fimbriae* are similar to pili but shorter, and help cells to adhere to surfaces such as animal cells, for food and protection.



5.5 Prokaryotic Flagella (A) Flagella contribute to the movement and adhesion of prokaryotic cells. (B) Complex protein ring structures anchored in the plasma membrane form a motor unit that rotates the flagellum and propels the cell.

CYTOSKELETON Some prokaryotes, especially rod-shaped bacteria, have a helical network of filamentous structures that extend down the length of the cell just inside the plasma membrane. The proteins that make up this structure are similar in amino acid sequence to actin in eukaryotic cells. Since actin is part of the cytoskeleton in eukaryotes (see Section 5.3), it has been suggested that the helical filaments in prokaryotes play a role in maintaining the rod-like cell shape.

5.2 RECAP

Prokaryotic organisms can live on diverse energy sources and in extreme environments. Unlike eukaryotic cells, prokaryotic cells do not have extensive internal compartments.

- What structures are present in all prokaryotic cells? See p. 82 and Figure 5.4
- Describe the structure and function of a specialized prokaryotic cell feature, such as the cell wall, capsule, flagellum, or pilus. See pp. 83–84 and Figure 5.5

As we mentioned earlier, the prokaryotic cell is one of two types of cell structure recognized in cell biology. The other is the eukaryotic cell. Eukaryotic cells, and multicellular eukaryotic organisms, are more structurally and functionally complex than prokaryotic cells.

5.3 What Features Characterize Eukaryotic Cells?

Eukaryotic cells generally have dimensions up to 10 times greater than those of prokaryotes; for example, the spherical yeast cell has a diameter of about 8 μm , in contrast to a typical bacterium with a diameter of 1 μm . Like prokaryotic cells, eukaryotic cells have a plasma membrane, cytoplasm, and ribosomes. But as you learned earlier in this chapter, eukaryotic cells also have compartments within the cytoplasm whose interiors are separated from the cytosol by membranes.

Compartmentalization is the key to eukaryotic cell function

The membranous compartments of eukaryotic cells are called **organelles**. Each type of organelle has a specific role in its particular cell. Some of the organelles have been characterized as factories that make specific products. Others are like power plants that take in energy in one form and convert it into a more useful form. These functional roles are defined by the chemical reactions each organelle can carry out:

- The *nucleus* contains most of the cell's genetic material (DNA). The replication of the genetic material and the first steps in expressing genetic information take place in the nucleus.
- The *mitochondrion* is a power plant and industrial park, where energy stored in the bonds of carbohydrates and

fatty acids is converted into a form that is more useful to the cell (ATP; see Section 9.1).

- The *endoplasmic reticulum* and *Golgi apparatus* are compartments in which some proteins synthesized by the ribosomes are packaged and sent to appropriate locations in the cell.
- *Lysosomes* and *vacuoles* are cellular digestive systems in which large molecules are hydrolyzed into usable monomers.
- *Chloroplasts* (found in only some cells) perform photosynthesis.

The membrane surrounding each organelle has two essential roles. First, it keeps the organelle's molecules away from other molecules in the cell, to prevent inappropriate reactions. Second, it acts as a traffic regulator, letting important raw materials into the organelle and releasing its products to the cytoplasm. In some organelles, the membrane also has proteins that have functional roles in chemical reactions that occur at the organelle surface.

There are a number of other structures in eukaryotic cells that have specialized functions, but are not generally called organelles because they lack membranes:

- Ribosomes, where protein synthesis takes place
- The cytoskeleton, composed of several types of protein-based filaments, which has both structural and functional roles
- The extracellular matrix, which also has structural and functional roles

The evolution of compartments was an important development that enabled eukaryotic cells to specialize, forming the organs and tissues of complex multicellular organisms.

Organelles can be studied by microscopy or isolated for chemical analysis

Cell organelles and structures were first detected by light and then by electron microscopy. The functions of the organelles could sometimes be inferred by observations and experiments, leading, for example, to the hypothesis (later confirmed) that the nucleus contained the genetic material. Later, the use of stains targeted to specific macromolecules allowed cell biologists to determine the chemical compositions of organelles (see Figure 5.17, which shows a single cell stained for three different proteins).

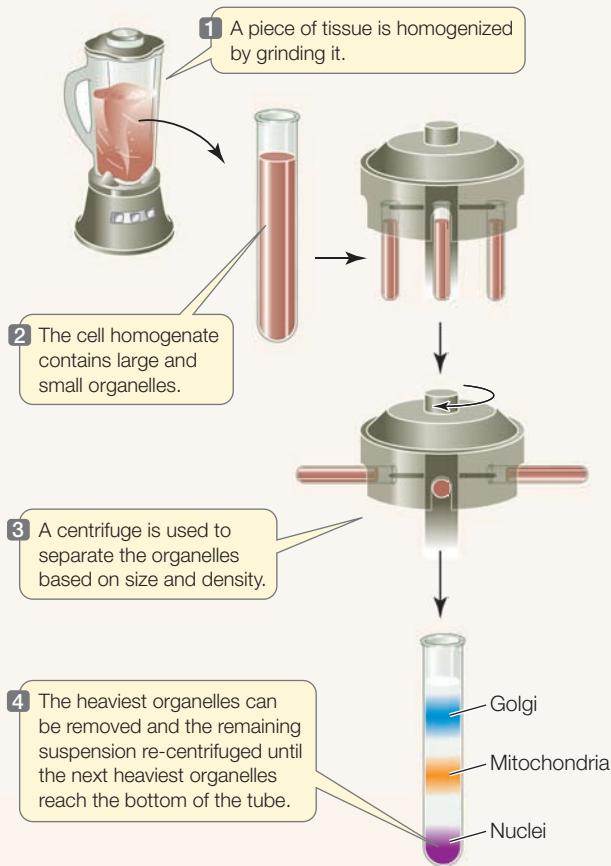
Another way to analyze cells is to take them apart in a process called cell fractionation. This process permits cell organelles and other cytoplasmic structures to be separated from each other and examined using chemical methods. Cell fractionation begins with the destruction of the plasma membrane, which allows the cytoplasmic components to flow out into a test tube. The various organelles can then be separated from one another on the basis of size or density (Figure 5.6). Biochemical analyses can then be done on the isolated organelles.

Microscopy and cell fractionation have complemented each other, giving us a more complete picture of the composition and function of each organelle and structure.

TOOLS FOR INVESTIGATING LIFE

5.6 Cell Fractionation

Organelles can be separated from one another after cells are broken open and their contents suspended in an aqueous medium. The medium is placed in a tube and spun in a centrifuge, which rotates about an axis at high speed. Centrifugal forces cause particles to sediment at the bottom of the tube where they may be collected for biochemical study. Heavier particles sediment at lower speeds than do lighter particles. By adjusting the speed of centrifugation, cellular organelles and even large particles like ribosomes can be separated and partially purified.



Microscopy of plant and animal cells has revealed that many of the organelles are similar in appearance in each cell type (**Figure 5.7**). By comparing the illustrations in Figure 5.7 and Figure 5.4 you can see some of the prominent differences between eukaryotic cells and prokaryotic cells.

Ribosomes are factories for protein synthesis

The ribosomes of prokaryotes and eukaryotes are similar in that both types consist of two different-sized subunits. Eukaryotic ribosomes are somewhat larger than those of prokaryotes, but the structure of prokaryotic ribosomes is better understood. Chemically, ribosomes consist of a special type of RNA called ribosomal RNA (rRNA). Ribosomes also contain more than 50

different protein molecules, which are noncovalently bound to the rRNA.

In prokaryotic cells, ribosomes float freely in the cytoplasm. In eukaryotic cells they are found in two places: in the cytoplasm, where they may be free or attached to the surface of the endoplasmic reticulum (a membrane-bound organelle, see below), and inside mitochondria and chloroplasts. In each of these locations, the ribosomes are molecular factories where proteins are synthesized with their amino acid sequences specified by nucleic acids. Although they seem small in comparison to the cells that contain them, by molecular standards ribosomes are huge complexes (about 25 nm in diameter), made up of several dozen different molecules.

The nucleus contains most of the genetic information

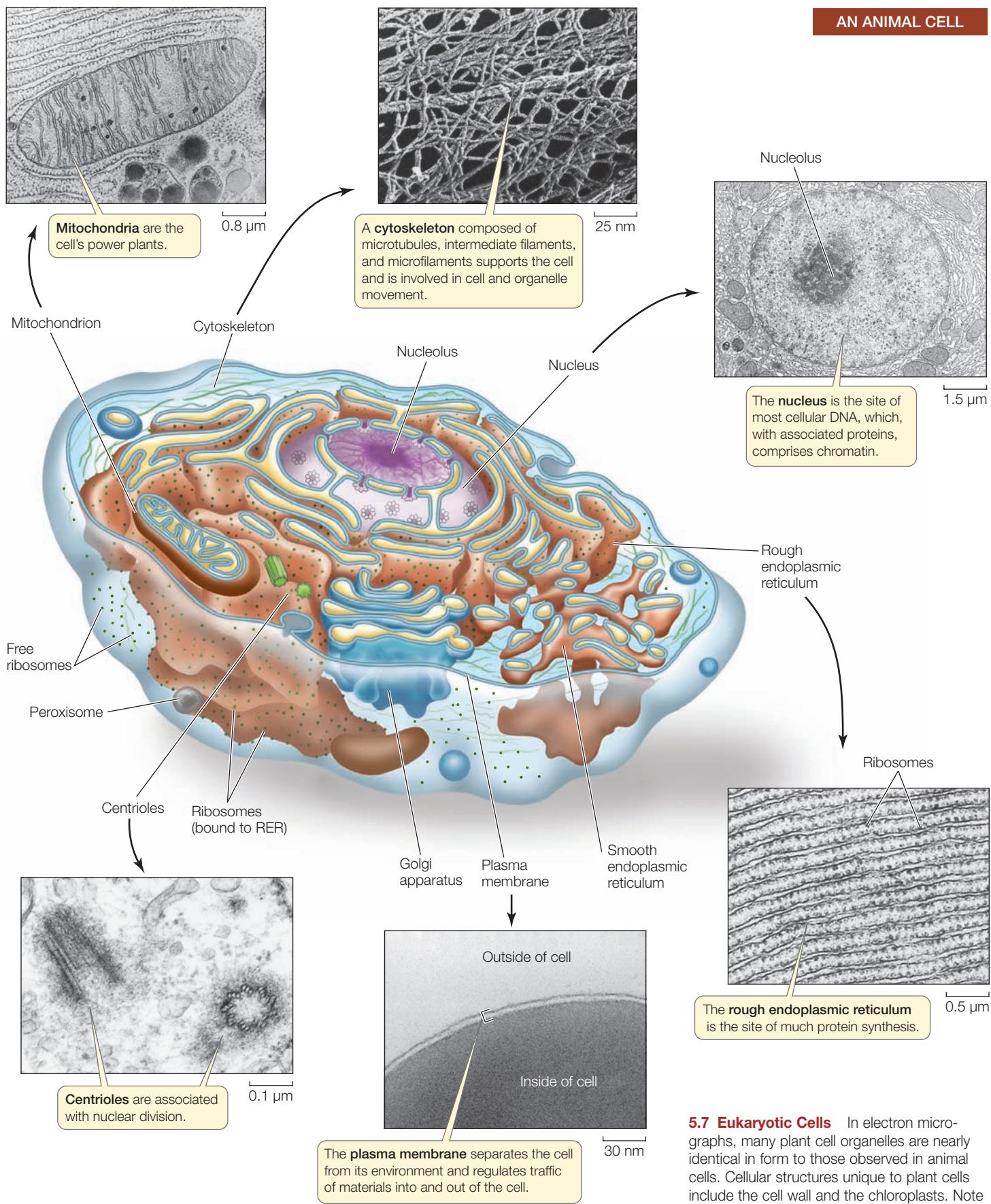
Organisms depend on accurate information—internal signals, environmental cues, and stored instructions—in order to respond appropriately to changing conditions, to maintain a constant internal environment, and to reproduce. In the cell, hereditary information is stored in the sequence of nucleotides in DNA molecules. Most of the DNA in eukaryotic cells resides in the nucleus (see Figure 5.7). Information encoded in the DNA is *translated* into proteins at the ribosomes. This process is described in detail in Chapter 14.

Most cells have a single nucleus, which is usually the largest organelle (**Figure 5.8**). The nucleus of a typical animal cell is approximately 5 μm in diameter—substantially larger than most prokaryotic cells. The nucleus has several functions in the cell:

- It is the location of the DNA and the site of DNA replication.
- It is the site where gene transcription is turned on or off.
- A region within the nucleus, the **nucleolus**, is where ribosomes begin to be assembled from RNA and proteins.

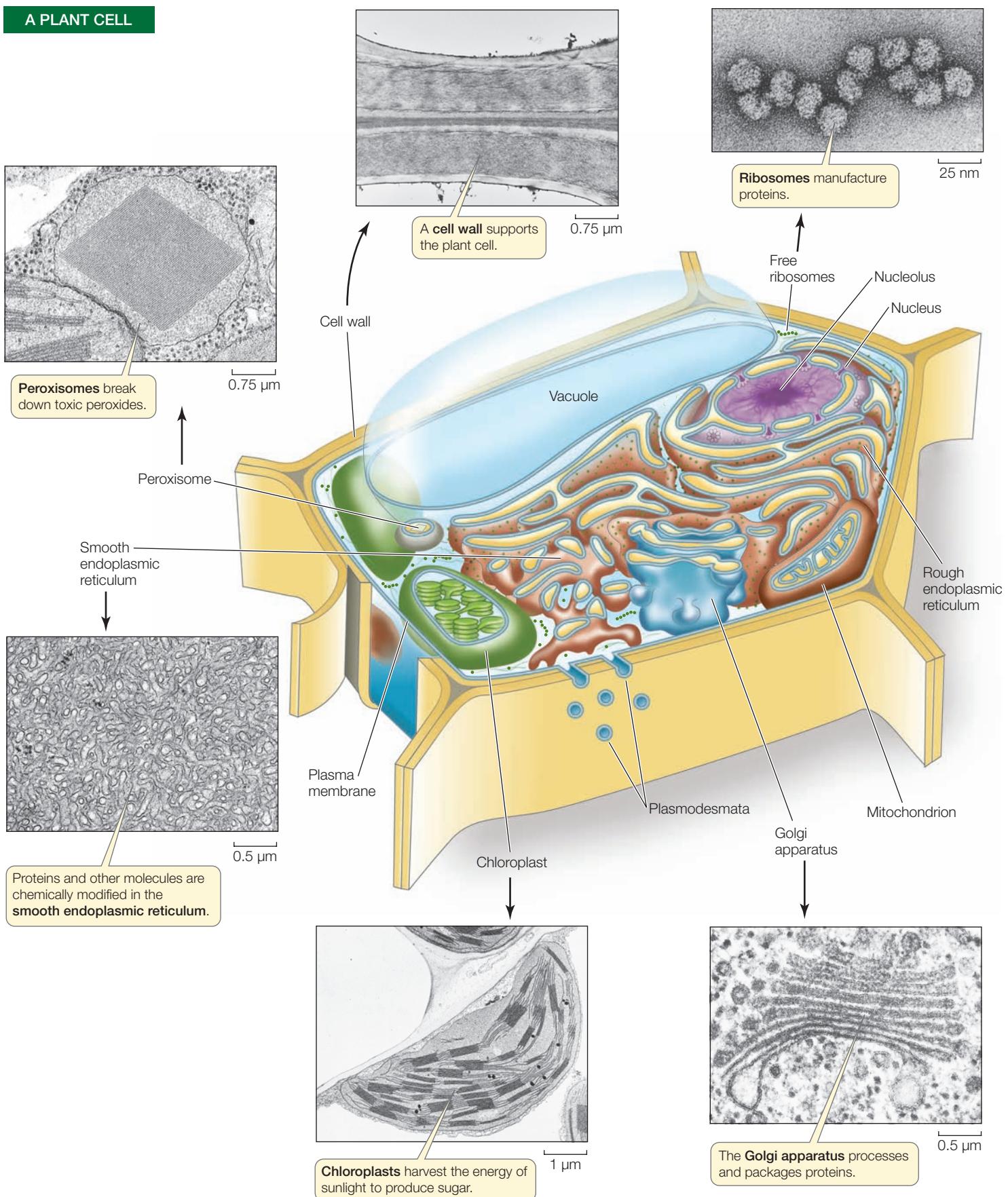
The nucleus is surrounded by two membranes, which together form the *nuclear envelope* (see Figure 5.8). This structure separates the genetic material from the cytoplasm. Functionally, it separates DNA transcription (which occurs in the nucleus) from translation (in the cytoplasm) (see Figure 4.5). The two membranes of the nuclear envelope are perforated by thousands of nuclear pores, each measuring approximately 9 nm in diameter, which connect the interior of the nucleus with the cytoplasm (see Figure 5.8). The pores regulate the traffic between these two cellular compartments by allowing some molecules to enter the nucleus and blocking others. This allows the nucleus to regulate the information-processing functions.

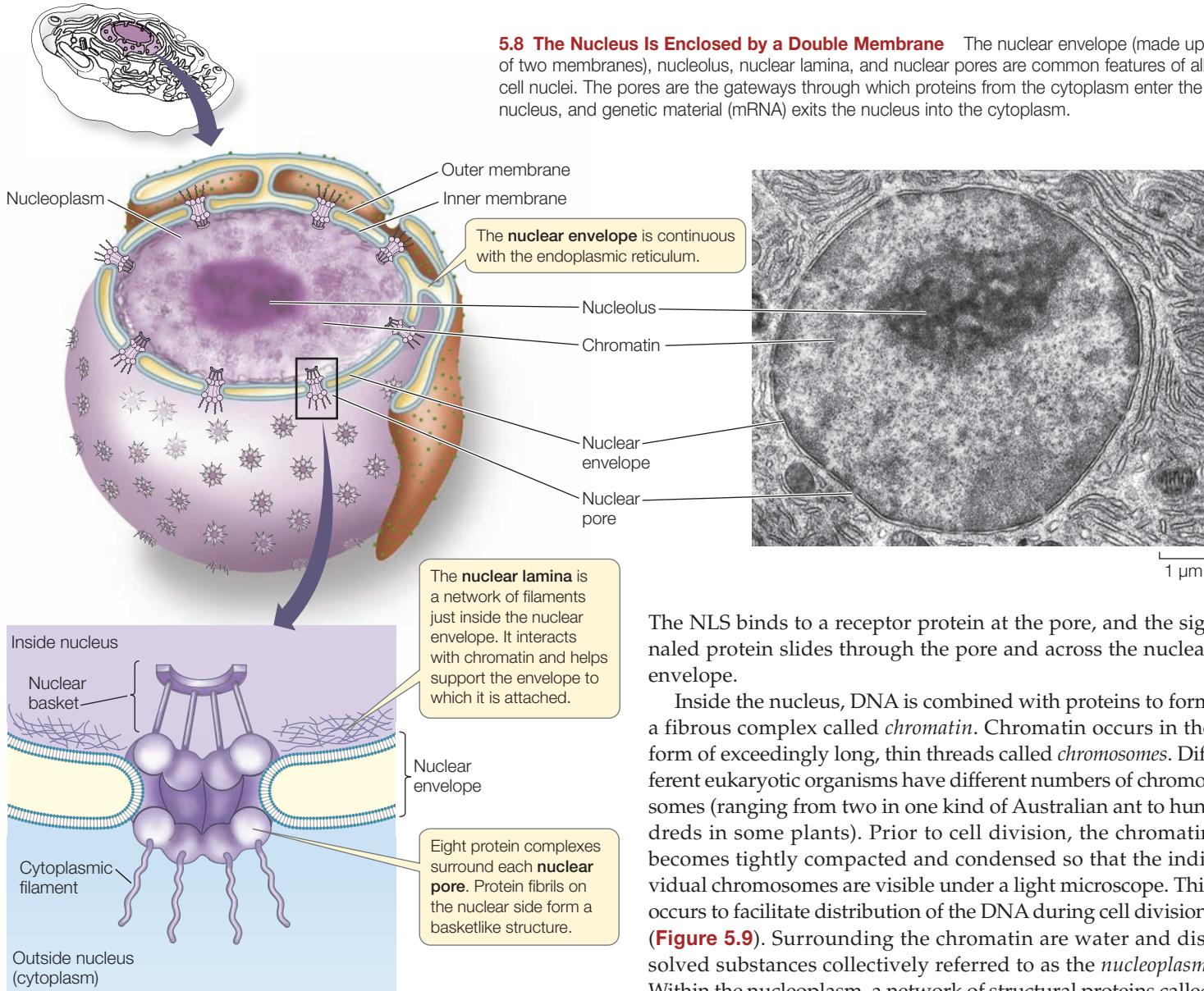
At the nuclear pore, small substances, including ions and other molecules with molecular weights of less than 10,000 dal-



ANIMAL CELLS

5.7 Eukaryotic Cells In electron micrographs, many plant cell organelles are nearly identical in form to those observed in animal cells. Cellular structures unique to plant cells include the cell wall and the chloroplasts. Note that the images are two-dimensional “slices,” while cells are three-dimensional structures.





tons, freely diffuse through the pore. Larger molecules, such as many proteins that are made in the cytoplasm and imported into the nucleus, cannot get through without a certain short sequence of amino acids that is part of the protein. We know that this sequence is the *nuclear localization signal* (NLS) from several lines of evidence (see also Figure 14.20):

- The NLS occurs in most proteins targeted to the nucleus, but not in proteins that remain in the cytoplasm.
- If the NLS is removed from a protein, the protein stays in the cytoplasm.
- If the NLS is added to a protein that normally stays in the cytoplasm, that protein moves into the nucleus.
- Some viruses have an NLS that allows them to enter the nucleus; viruses without the signal sequence do not enter the nucleus as virus particles.

5.8 The Nucleus Is Enclosed by a Double Membrane The nuclear envelope (made up of two membranes), nucleolus, nuclear lamina, and nuclear pores are common features of all cell nuclei. The pores are the gateways through which proteins from the cytoplasm enter the nucleus, and genetic material (mRNA) exits the nucleus into the cytoplasm.

The NLS binds to a receptor protein at the pore, and the signaled protein slides through the pore and across the nuclear envelope.

Inside the nucleus, DNA is combined with proteins to form a fibrous complex called *chromatin*. Chromatin occurs in the form of exceedingly long, thin threads called *chromosomes*. Different eukaryotic organisms have different numbers of chromosomes (ranging from two in one kind of Australian ant to hundreds in some plants). Prior to cell division, the chromatin becomes tightly compacted and condensed so that the individual chromosomes are visible under a light microscope. This occurs to facilitate distribution of the DNA during cell division. (Figure 5.9). Surrounding the chromatin are water and dissolved substances collectively referred to as the *nucleoplasm*. Within the nucleoplasm, a network of structural proteins called the *nuclear matrix* helps organize the chromatin.

At the interior periphery of the nucleus, the chromatin is attached to a protein meshwork, called the *nuclear lamina*, which is formed by the polymerization of proteins called lamins into long thin structures called intermediate filaments. The nuclear lamina maintains the shape of the nucleus by its attachment to both the chromatin and the nuclear envelope. There is some evidence that the nuclear lamina may be involved with human aging. As people age, the nuclear lamina begins to disintegrate and in the process the structural integrity of the nucleus declines. In people with the rare disease called progeria, this decline begins very early in life and their aging is accelerated.

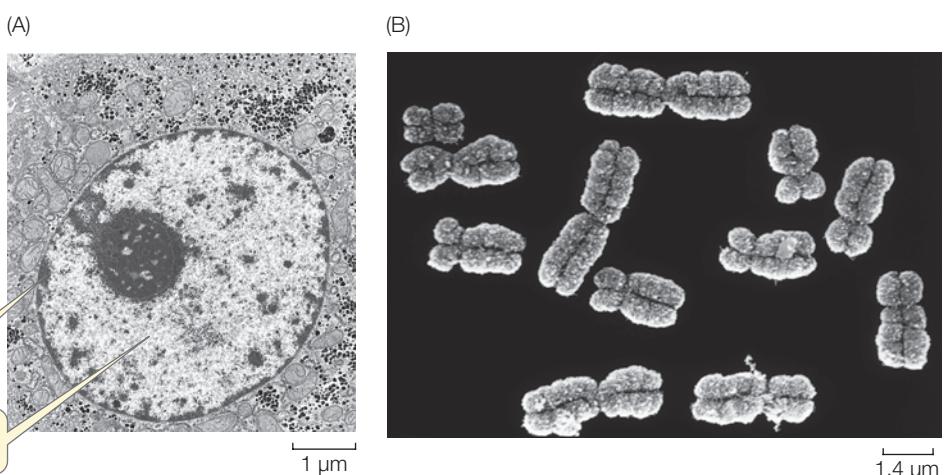
During most of a cell's life cycle, the nuclear envelope is a stable structure. When the cell reproduces, however, the nuclear envelope breaks down into small, membrane-bound droplets, called *vesicles*, containing pore complexes. The envelope reforms after the replicated DNA has been distributed to the daughter cells (see Section 11.3).

At certain sites, the outer membrane of the nuclear envelope folds outward into the cytoplasm and is continuous with the

5.9 Chromatin and Chromosomes

(A) When a cell is not dividing, the nuclear DNA is aggregated with proteins to form chromatin, which is dispersed throughout the nucleus. This two-dimensional image was made using a transmission electron microscope. (B) The chromosomes in dividing cells become highly condensed. This three-dimensional image of isolated metaphase chromosomes was produced by a scanning electron microscope.

Dense chromatin (dark) near the nuclear envelope is attached to the nuclear lamina.
Diffuse chromatin (light) is in the nucleoplasm.



membrane of another organelle, the endoplasmic reticulum, which we will discuss next.

The endomembrane system is a group of interrelated organelles

Much of the volume of some eukaryotic cells is taken up by an extensive **endomembrane system**. This is an interconnected system of membrane-enclosed compartments that are sometimes flattened into sheets and sometimes have other characteristic shapes (see Figure 5.7). The endomembrane system includes the plasma membrane, nuclear envelope, endoplasmic reticulum, Golgi apparatus, and lysosomes, which are derived from the Golgi. Tiny, membrane-surrounded droplets called vesicles shuttle substances between the various components of the endomembrane system (Figure 5.10). In drawings and electron microscope pictures this system appears static, fixed in space and time. But these depictions are just snapshots; in the living cell, membranes and the materials they contain are in constant motion. Membrane components have been observed to shift from one organelle to another within the endomembrane system. Thus, all these membranes must be functionally related.

ENDOPLASMIC RETICULUM Electron micrographs of eukaryotic cells reveal networks of interconnected membranes branching throughout the cytoplasm, forming tubes and flattened sacs. These membranes are collectively called the **endoplasmic reticulum**, or **ER**. The interior compartment of the ER, referred to as the lumen, is separate and distinct from the surrounding cytoplasm (see Figure 5.10). The ER can enclose up to 10 percent of the interior volume of the cell, and its foldings result in a surface area many times greater than that of the plasma membrane. There are two types of endoplasmic reticulum, the so-called rough and smooth.

Rough endoplasmic reticulum (RER) is called “rough” because of the many ribosomes attached to the outer surface of the membrane, giving it a “rough” appearance in electron microscopy (see Figure 5.7). The attached ribosomes are actively involved in protein synthesis, but that is not the entire story:

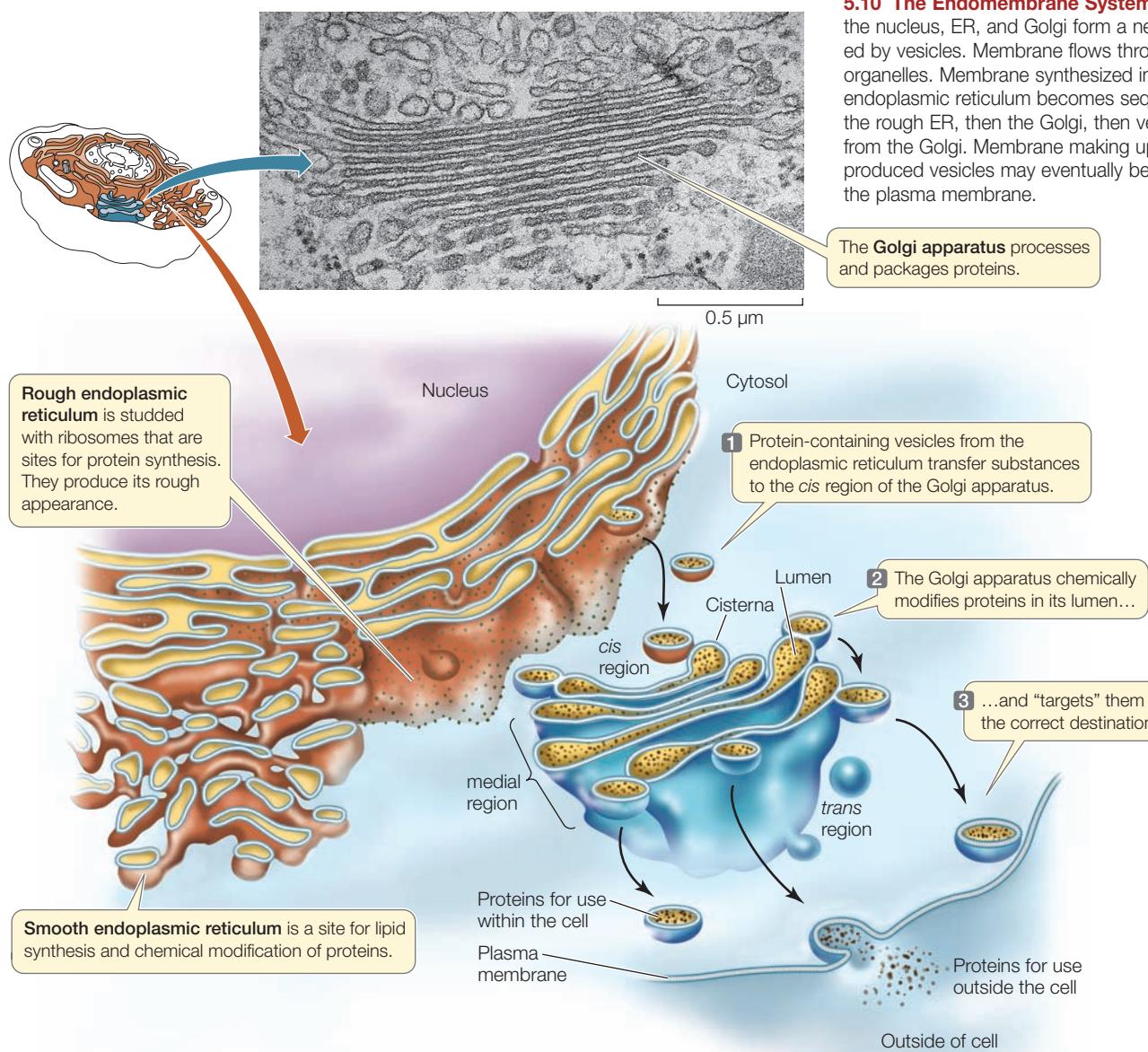
- The RER receives into its lumen certain newly synthesized proteins, segregating them away from the cytoplasm. The RER also participates in transporting these proteins to other locations in the cell.
- While inside the RER, proteins can be chemically modified to alter their functions and to chemically ‘tag’ them for delivery to specific cellular destinations.
- Proteins are shipped to cellular destinations enclosed within vesicles that pinch off from the ER.
- Most membrane-bound proteins are made in the RER.

A protein enters the lumen of the RER through a pore as it is synthesized. As with a protein passing through a nuclear pore, this is accomplished via a sequence of amino acids on the protein, which acts as a RER localization signal (see Section 14.6). Once in the lumen of the RER, proteins undergo several changes, including the formation of disulfide bridges and folding into their tertiary structures (see Figure 3.7).

Some proteins are covalently linked to carbohydrate groups in the RER, thus becoming glycoproteins. In the case of proteins directed to the lysosomes, the carbohydrate groups are part of an “addressing” system that ensures that the right proteins are directed to those organelles. This addressing system is very important because the enzymes within the lysosomes are some of the most destructive the cell makes. Were they not properly addressed and contained, they could destroy the cell.

The **smooth endoplasmic reticulum (SER)** lacks ribosomes and is more tubular (and less like flattened sacs) than the RER, but it shows continuity with portions of the RER (see Figure 5.10). Within the lumen of the SER, some proteins that have been synthesized on the RER are chemically modified. In addition, the SER has three other important roles:

- It is responsible for the chemical modification of small molecules taken in by the cell, including drugs and pesticides. These modifications make the targeted molecules more polar, so they are more water-soluble and more easily removed.
- It is the site for glycogen degradation in animal cells. We discuss this important process in Chapter 9.
- It is the site for the synthesis of lipids and steroids.



Cells that synthesize a lot of protein for export are usually packed with RER. Examples include glandular cells that secrete digestive enzymes and white blood cells that secrete antibodies. In contrast, cells that carry out less protein synthesis (such as storage cells) contain less RER. Liver cells, which modify molecules (including toxins) that enter the body from the digestive system, have abundant SER.

GOLGI APPARATUS The **Golgi apparatus** (or Golgi complex), more often referred to merely as the Golgi, is another part of the diverse, dynamic, and extensive endomembrane system (see Figure 5.10). The exact appearance of the Golgi apparatus (named for its discoverer, Camillo Golgi) varies from species to species, but it almost always consists of two components: flattened membranous sacs called *cisternae* (singular *cisterna*) that are piled up like saucers, and small membrane-enclosed vesicles. The entire apparatus is about 1 μm long.

5.10 The Endomembrane System Membranes of the nucleus, ER, and Golgi form a network, connected by vesicles. Membrane flows through these organelles. Membrane synthesized in the smooth endoplasmic reticulum becomes sequentially part of the rough ER, then the Golgi, then vesicles formed from the Golgi. Membrane making up the Golgi-produced vesicles may eventually become part of the plasma membrane.

The **Golgi apparatus** processes and packages proteins.

The Golgi has several roles:

- When protein-containing vesicles from the RER fuse with the Golgi membranes, the Golgi receives the proteins and may further modify them.
- It concentrates, packages, and sorts proteins before they are sent to their cellular or extracellular destinations.
- It adds some carbohydrates to proteins and modifies others that were attached to proteins in the ER.
- It is where some polysaccharides for the plant cell wall are synthesized.

While there is a characteristic form for all Golgi, there are also variations in its size and appearance in different cell types. In the cells of plants, protists, fungi, and many invertebrate animals, the stacks of cisternae are individual units scattered throughout the cytoplasm. In vertebrate cells, a few such stacks usually form a single, larger, more complex Golgi apparatus.

The cisternae of the Golgi apparatus appear to have three functionally distinct regions: the *cis* region lies nearest to the nucleus or a patch of RER, the *trans* region lies closest to the plasma membrane, and the *medial* region lies in between (see Figure 5.10). (The terms *cis*, *trans*, and *medial* derive from Latin words meaning, respectively, “on the same side,” “on the opposite side,” and “in the middle.”) These three parts of the Golgi apparatus contain different enzymes and perform different functions.

The Golgi apparatus receives proteins from the ER, packages them, and sends them on their way. Since there is often no direct membrane continuity between the ER and Golgi apparatus, how does a protein get from one organelle to the other? The protein could simply leave the ER, travel across the cytoplasm, and enter the Golgi apparatus. But that would expose the protein to interactions with other molecules in the cytoplasm. On the other hand, segregation from the cytoplasm could be maintained if a piece of the ER could “bud off,” forming a membranous vesicle that contains the protein—and that is exactly what happens.

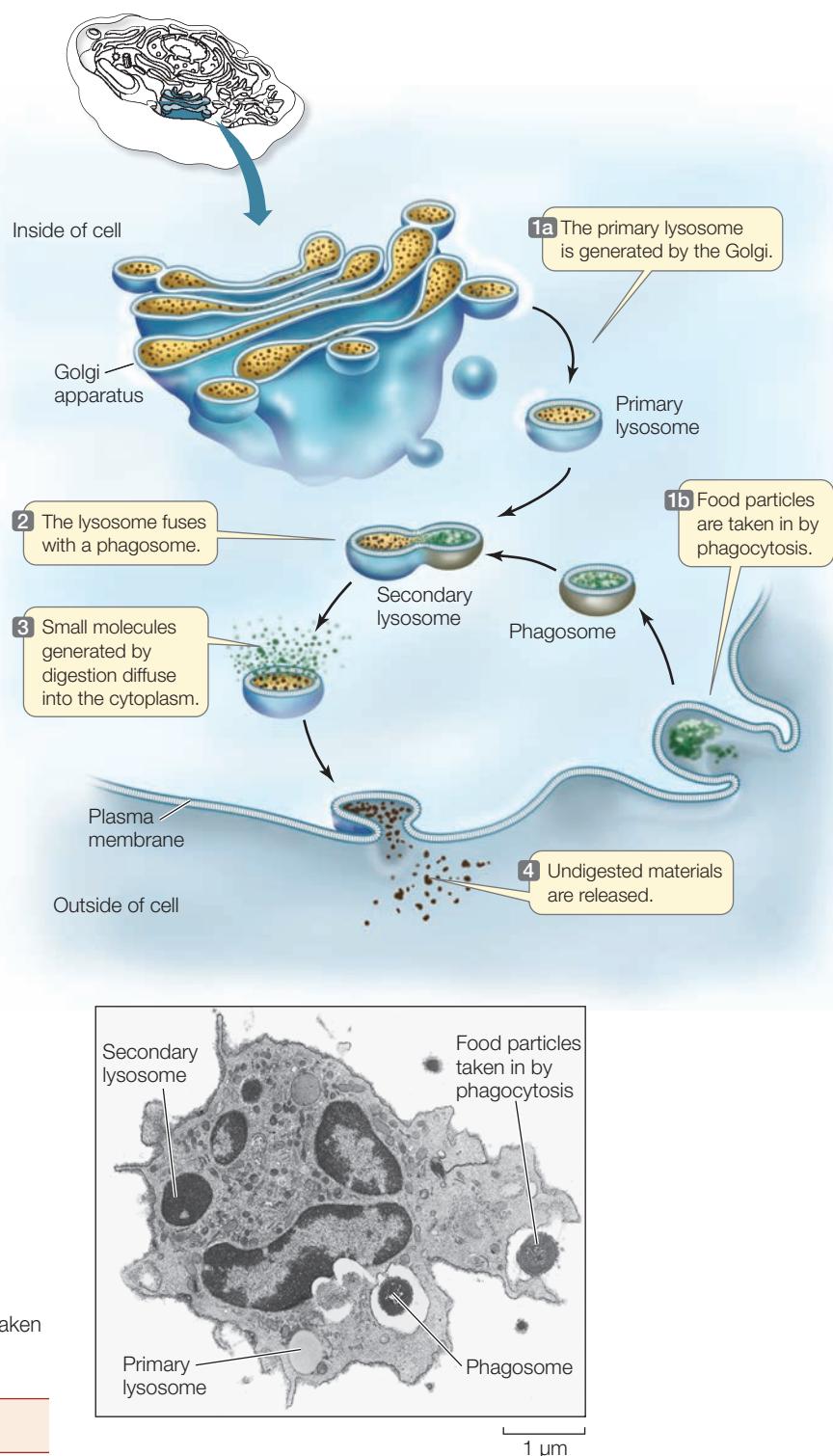
Proteins make the passage from the ER to the Golgi apparatus safely enclosed in vesicles. Once it arrives, a vesicle fuses with the *cis* membrane of the Golgi apparatus, releasing its cargo into the lumen of the Golgi cisterna. Other vesicles may move between the cisternae, transporting proteins, and it appears that some proteins move from one cisterna to the next through tiny channels. Vesicles budding off from the *trans* region carry their contents away from the Golgi apparatus. These vesicles go to the plasma membrane, or to another organelle in the endomembrane system called the lysosome.

LYSOSOMES The **primary lysosomes** originate from the Golgi apparatus. They contain digestive enzymes, and they are the sites where macromolecules—proteins, polysaccharides, nucleic acids, and lipids—are hydrolyzed into their monomers (see Figure 3.4). Lysosomes are about 1 μm in diameter; they are surrounded by a single membrane and have a featureless interior (Figure 5.11). There may be dozens of lysosomes in a cell, depending on its needs.

Lysosomes are sites for the breakdown of food, other cells, or foreign objects that are taken up by the cell. These materials get into the cell by a process called *phagocytosis* (*phago*, “eat”; *cytosis*, “cellular”). In this process, a pocket forms in the plasma membrane and then deepens and encloses material from outside

the cell. The pocket becomes a small vesicle called a phagosome, containing food or other material, which breaks free of the plasma membrane to move into the cytoplasm. The phagosome fuses with a primary lysosome to form a **secondary lysosome**, in which digestion occurs.

The effect of this fusion is rather like releasing hungry foxes into a chicken coop: the enzymes in the secondary lysosome quickly hydrolyze the food particles. These reactions are en-



5.11 Lysosomes Isolate Digestive Enzymes from the Cytoplasm Lysosomes are sites for the hydrolysis of material taken into the cell by phagocytosis.

hanced by the mild acidity of the lysosome's interior, where the pH is lower than in the surrounding cytoplasm. The products of digestion pass through the membrane of the lysosome, providing energy and raw materials for other cellular processes. The "used" secondary lysosome, now containing undigested particles, then moves to the plasma membrane, fuses with it, and releases the undigested contents to the environment.

Phagocytes are specialized cells that have an essential role in taking up and breaking down materials; they are found in nearly all animals and many protists. You will encounter them and their activities again at many places in this book, but at this point one example suffices: in the human liver and spleen, phagocytes digest approximately 10 billion aged or damaged blood cells each day! The digestion products are then used to make new cells to replace those that are digested.

Lysosomes are active even in cells that do not perform phagocytosis. Because cells are such dynamic systems, some cell components are frequently destroyed and replaced by new ones. The programmed destruction of cell components is called *autophagy*, and lysosomes are where the cell breaks down its own materials. With the proper signal, lysosomes can engulf entire organelles, hydrolyzing their constituents.

How important is autophagy? An entire class of human diseases called lysosomal storage diseases occur when lysosomes fail to digest internal components; these diseases are invariably very harmful or fatal. An example is Tay-Sachs disease, in which a particular lipid called a ganglioside is not broken down in lysosomes and instead accumulates in brain cells. In the most common form of this disease, a baby starts exhibiting neurological symptoms and becomes blind, deaf, and unable to swallow after six months of age. Death occurs before age 4.

Plant cells do not appear to contain lysosomes, but the central vacuole of a plant cell (which we will describe below) may function in an equivalent capacity because it, like lysosomes, contains many digestive enzymes.

Some organelles transform energy

All living things require external sources of energy. The energy from such sources must be transformed so that it can be used by cells. A cell requires energy to make the molecules it needs for activities such as growth, reproduction, responsiveness, and movement. Energy is transformed from one form to another in mitochondria (found in all eukaryotic cells) and in chloroplasts (found in eukaryotic cells that harvest energy from sunlight). In contrast, energy transformations in prokaryotic cells are associated with enzymes attached to the inner surface of the plasma membrane or to extensions of the plasma membrane that protrude into the cytoplasm.

MITOCHONDRIA In eukaryotic cells, the breakdown of fuel molecules such as glucose begins in the cytosol. The molecules that result from this partial degradation enter the **mitochondria** (singular *mitochondrion*), whose primary function is to convert the chemical energy of those fuel molecules into a form that the cell can use, namely the energy-rich molecule ATP (adenosine triphosphate) (see Section 8.2). The production of ATP in the mi-

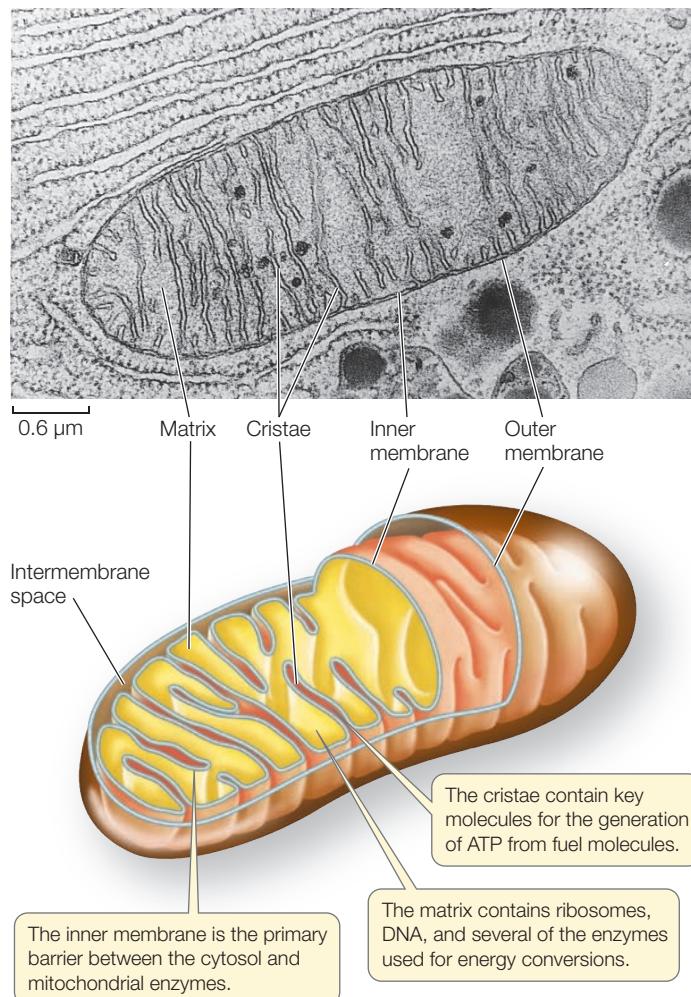
tochondria, using fuel molecules and molecular oxygen (O_2), is called *cellular respiration*.

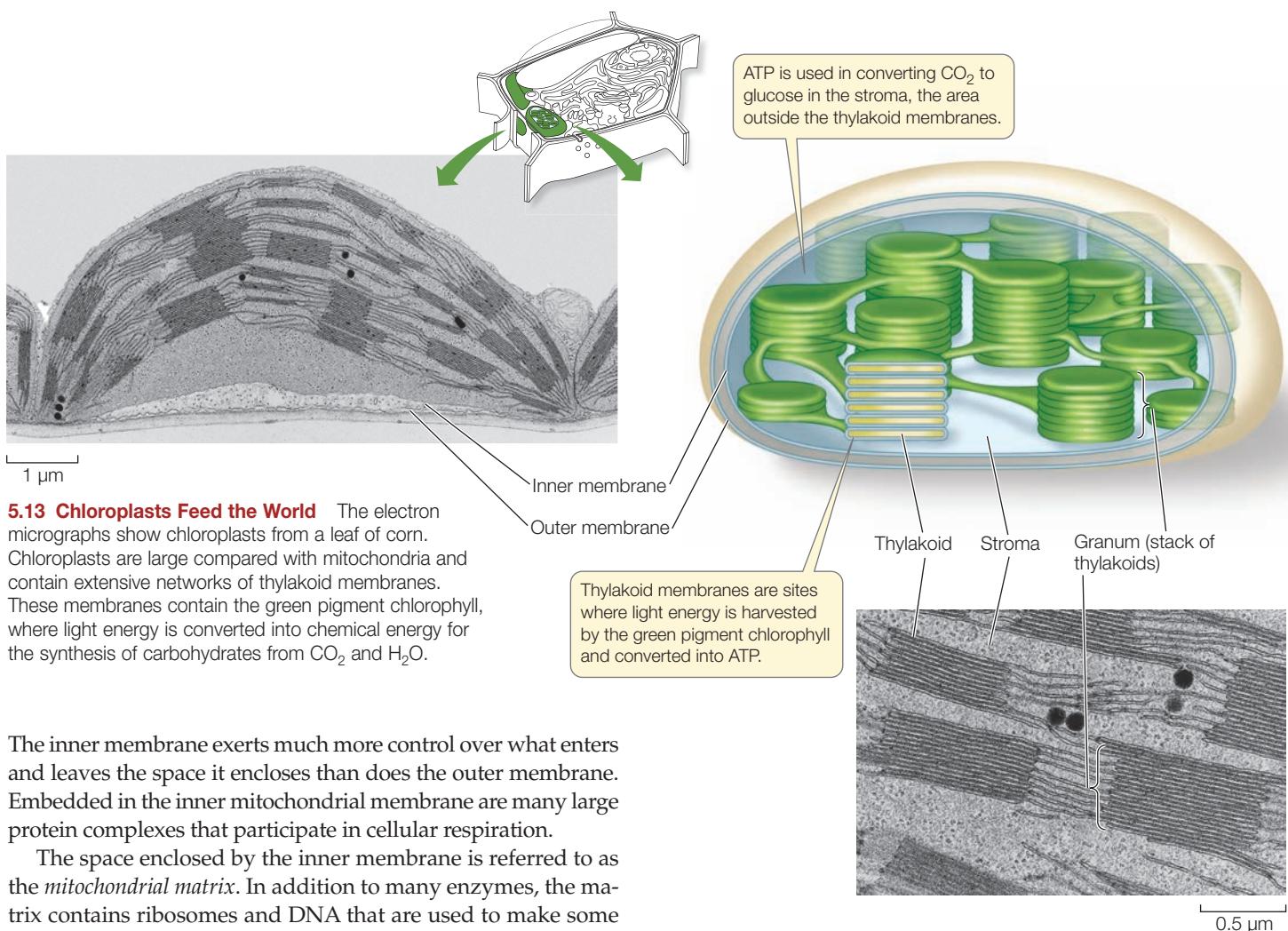
Typical mitochondria are somewhat less than 1.5 μm in diameter and 2–8 μm in length—about the size of many bacteria. They can divide independently of the central nucleus. The number of mitochondria per cell ranges from one gigantic organelle in some unicellular protists to a few hundred thousand in large egg cells. An average human liver cell contains more than a thousand mitochondria. Cells that are active in movement and growth require the most chemical energy, and these tend to have the most mitochondria per unit of volume.

Mitochondria have two membranes. The outer membrane is smooth and protective, and it offers little resistance to the movement of substances into and out of the organelle. Immediately inside the outer membrane is an inner membrane, which folds inward in many places, and thus has a surface area much greater than that of the outer membrane (Figure 5.12). The folds tend to be quite regular, giving rise to shelf-like structures called *cristae*.

5.12 A Mitochondrion Converts Energy from Fuel Molecules into ATP

The electron micrograph is a two-dimensional slice through a three-dimensional organelle. As the drawing emphasizes, the cristae are extensions of the inner mitochondrial membrane.



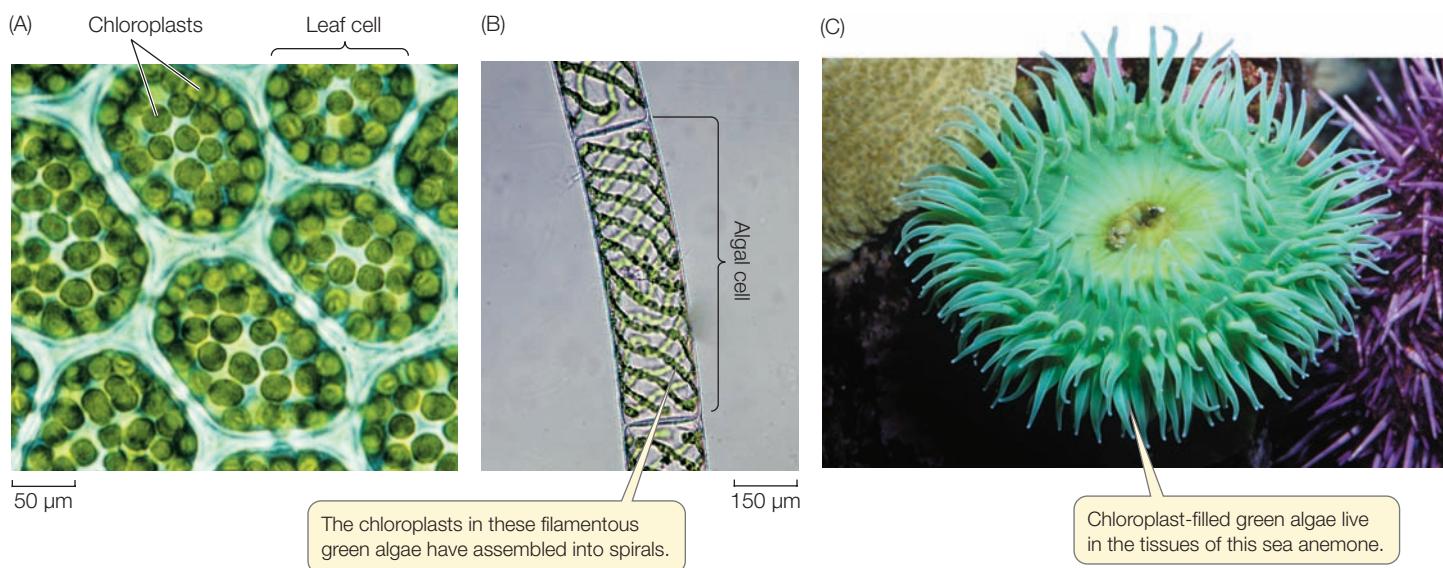


grana), consists of a series of flat, closely packed, circular compartments called **thylakoids** (see Figure 5.13). Thylakoid lipids are distinctive: only 10 percent are phospholipids, while the rest are galactose-substituted diglycerides and sulfolipids. Because of the abundance of chloroplasts, these are the most abundant lipids in the biosphere.

In addition to lipids and proteins, the membranes of the thylakoids contain chlorophyll and other pigments that harvest light energy for photosynthesis (we see how they do this in Section 10.2). The thylakoids of one grana may be connected to those of other grana, making the interior of the chloroplast a highly developed network of membranes, much like the ER.

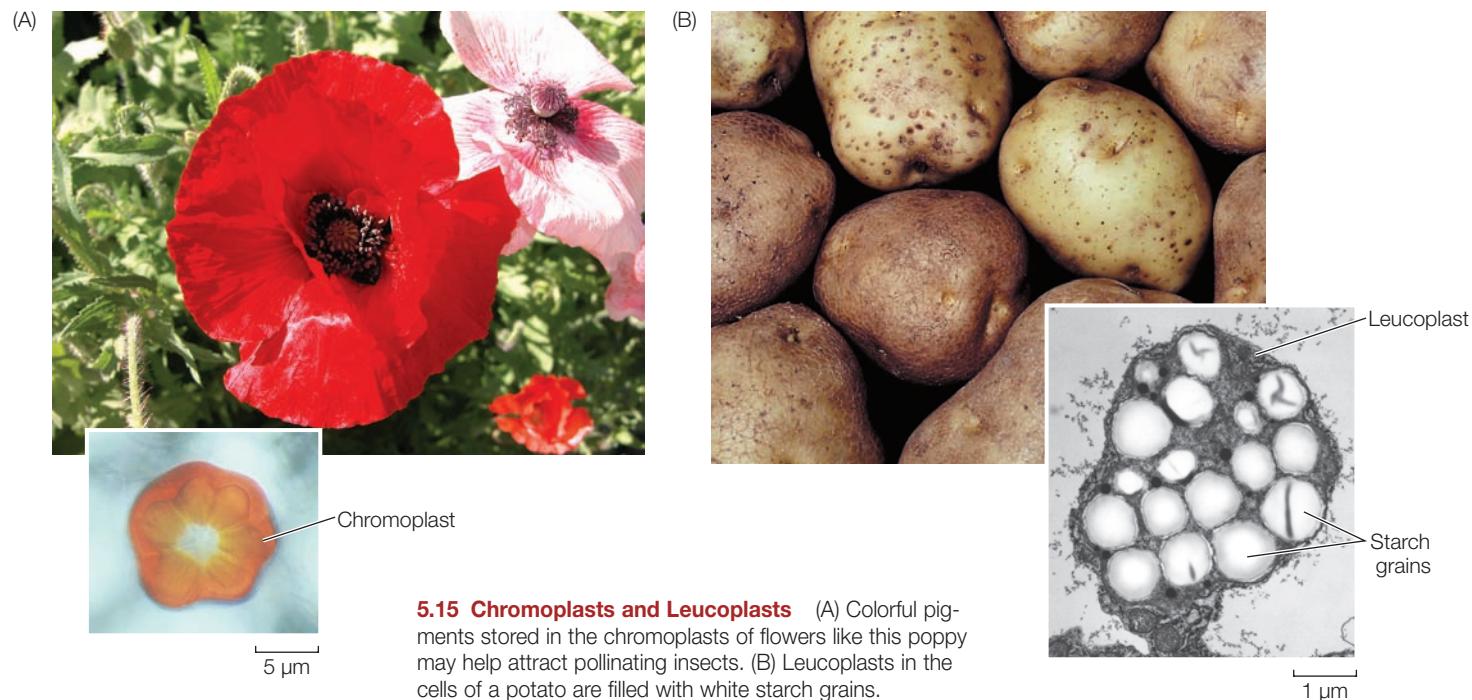
The fluid in which the grana are suspended is called the *stroma*. Like the mitochondrial matrix, the chloroplast stroma contains ribosomes and DNA, which are used to synthesize some, but not all, of the proteins that make up the chloroplast.

Animal cells typically do not contain chloroplasts, but some do contain functional photosynthetic organisms. The green color of some corals and sea anemones comes from chloroplasts in algae that live within those animals (see Figure 5.14C). The animals derive some of their nutrition from the photosynthesis that their chloroplast-containing “guests” carry out. Such an intimate relationship between two different organisms is called **symbiosis**.



5.14 Chloroplasts Are Everywhere (A) In green plants, chloroplasts are concentrated in the leaf cells. (B) Green algae are photosynthetic and filled with chloroplasts. (C) No animal species produces its own chloroplasts, but this sea anemone (an animal) is nourished by the chloroplasts of unicellular green algae living within its tissues, in what is termed a symbiotic relationship.

Other types of plastids such as *chromoplasts* and *leucoplasts* have functions different from those of chloroplasts (Figure 5.15). Chromoplasts make and store red, yellow, and orange pigments, especially in flowers and fruits. Leucoplasts are storage organelles that do not contain pigments. An amyloplast is a leucoplast that stores starch.

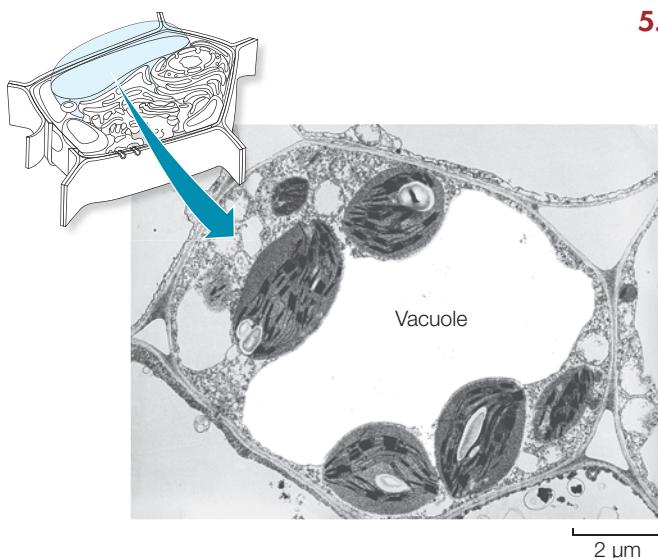


5.15 Chromoplasts and Leucoplasts (A) Colorful pigments stored in the chromoplasts of flowers like this poppy may help attract pollinating insects. (B) Leucoplasts in the cells of a potato are filled with white starch grains.

There are several other membrane-enclosed organelles

There are several other organelles whose boundary membranes separate their specialized chemical reactions and contents from the cytoplasm: peroxisomes, glyoxysomes, and vacuoles, including contractile vacuoles.

Peroxisomes are organelles that accumulate toxic peroxides, such as hydrogen peroxide (H_2O_2), that occur as byproducts of some biochemical reactions. These peroxides can be safely broken down inside the peroxisomes without mixing with other parts of the cell. Peroxisomes are small organelles, about 0.2 to 1.7 µm in diameter. They have a single membrane and a granular interior containing specialized enzymes. Peroxisomes are found in at least some of the cells of almost every eukaryotic species.



5.16 Vacuoles in Plant Cells Are Usually Large The large central vacuole in this cell is typical of mature plant cells. Smaller vacuoles are visible toward each end of the cell.

Glyoxysomes are similar to peroxisomes and are found only in plants. They are most abundant in young plants, and are the locations where stored lipids are converted into carbohydrates for transport to growing cells.

Vacuoles occur in many eukaryotic cells, but particularly those of plants and protists. Plant vacuoles (Figure 5.16) have several functions:

- **Storage:** Plant cells produce a number of toxic by-products and waste products, many of which are simply stored within vacuoles. Because they are poisonous or distasteful, these stored materials deter some animals from eating the plants, and may thus contribute to plant defenses and survival.
- **Structure:** In many plant cells, enormous vacuoles take up more than 90 percent of the cell volume and grow as the cell grows. The presence of dissolved substances in the vacuole causes water to enter it from the cytoplasm, making the vacuole swell like a balloon. The plant cell does not swell when the vacuole fills with water, since it has a rigid cell wall. Instead, it stiffens from the increase in water pressure (called turgor), which supports the plant (see Figure 6.10).
- **Reproduction:** Some pigments (especially blue and pink ones) in the petals and fruits of flowering plants are contained in vacuoles. These pigments—the anthocyanins—are visual cues that help attract the animals that assist in pollination or seed dispersal.
- **Digestion:** In some plants, vacuoles in seeds contain enzymes that hydrolyze stored seed proteins into monomers that the developing plant embryo can use as food.

Contractile vacuoles are found in many freshwater protists. Their function is to get rid of the excess water that rushes into the cell because of the imbalance in solute concentration between the interior of the cell and its freshwater environment. The contractile vacuole enlarges as water enters, then abruptly contracts, forcing the water out of the cell through a special pore structure.

So far, we have discussed numerous membrane-enclosed organelles. Now we turn to a group of cytoplasmic structures without membranes.

The cytoskeleton is important in cell structure and movement

From the earliest observations, light microscopy revealed distinctive shapes of cells that would sometimes change, and within cells rapid movements were observed. With the advent of electron microscopy, a new world of cellular substructure was revealed, including a meshwork of filaments inside cells. Experimentation showed that this **cytoskeleton** fills several important roles:

- It supports the cell and maintains its shape.
- It holds cell organelles in position within the cell.
- It moves organelles within the cell.
- It is involved with movements of the cytoplasm, called cytoplasmic streaming.
- It interacts with extracellular structures, helping to anchor the cell in place.

There are three components of the cytoskeleton: microfilaments (smallest diameter), intermediate filaments, and microtubules (largest diameter). These filaments have very different functions.

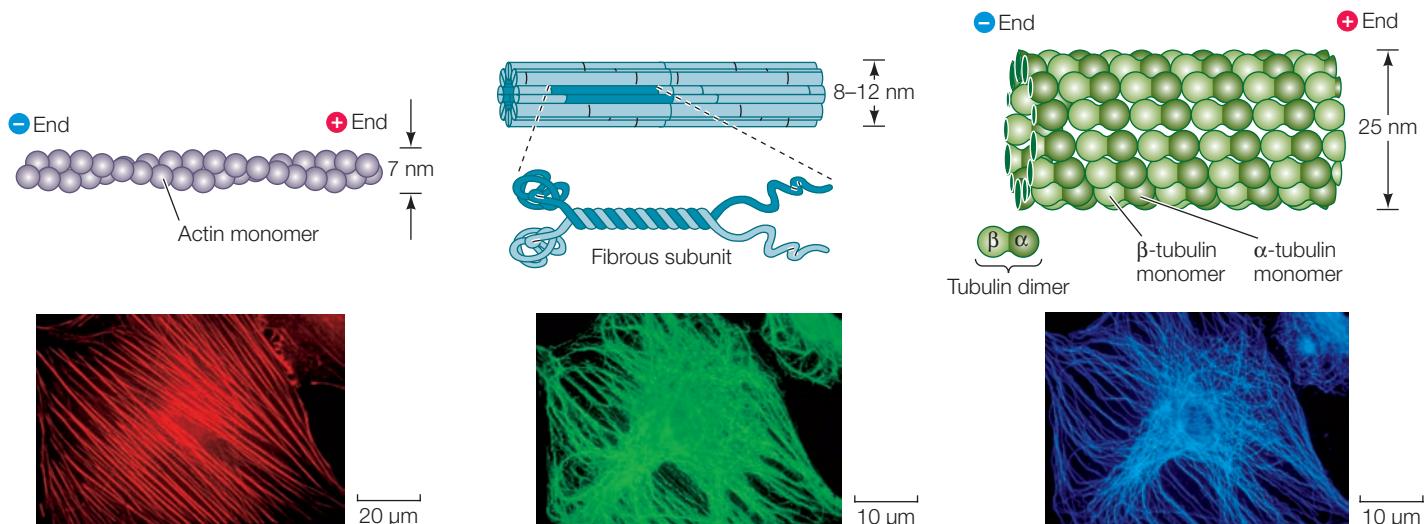
MICROFILAMENTS **Microfilaments** can exist as single filaments, in bundles, or in networks. They are about 7 nm in diameter and up to several micrometers long. Microfilaments have two major roles:

- They help the entire cell or parts of the cell to move.
- They determine and stabilize cell shape.

Microfilaments are assembled from *actin* monomers, a protein that exists in several forms and has many functions, especially in animals. The actin found in microfilaments (which are also known as actin filaments) has distinct ends designated “plus” and “minus.” These ends permit actin monomers to interact with one another to form long, double helical chains (Figure 5.17A). Within cells, the polymerization of actin into microfilaments is reversible, and the microfilaments can disappear from cells by breaking down into monomers of free actin. Special actin-binding proteins mediate these events.

In the muscle cells of animals, actin filaments are associated with another protein, the “motor protein” *myosin*, and the interactions of these two proteins account for the contraction of muscles (described in Section 48.1). In non-muscle cells, actin filaments are associated with localized changes in cell shape. For example, microfilaments are involved in the flowing movement of the cytoplasm called cytoplasmic streaming, in amoeboid movement, and in the “pinching” contractions that divide an animal cell into two daughter cells. Microfilaments are also involved in the formation of cellular extensions called pseudopodia (*pseudo*, “false”; *podia*, “feet”) that enable some cells to move (Figure 5.18). As you will see in Chapter 42, cells of the immune system must move toward other cells during the immune response.

In some cell types, microfilaments form a meshwork just inside the plasma membrane. Actin-binding proteins then cross-link the microfilaments to form a rigid net-like structure that supports the cell. For example, microfilaments support the tiny

**(A) Microfilaments**

Made up of strands of the protein actin; often interact with strands of other proteins.

(B) Intermediate filaments

Made up of fibrous proteins organized into tough, ropelike assemblages that stabilize a cell's structure and help maintain its shape.

(C) Microtubules

Long, hollow cylinders made up of many molecules of the protein tubulin. Tubulin consists of two subunits, α -tubulin and β -tubulin.

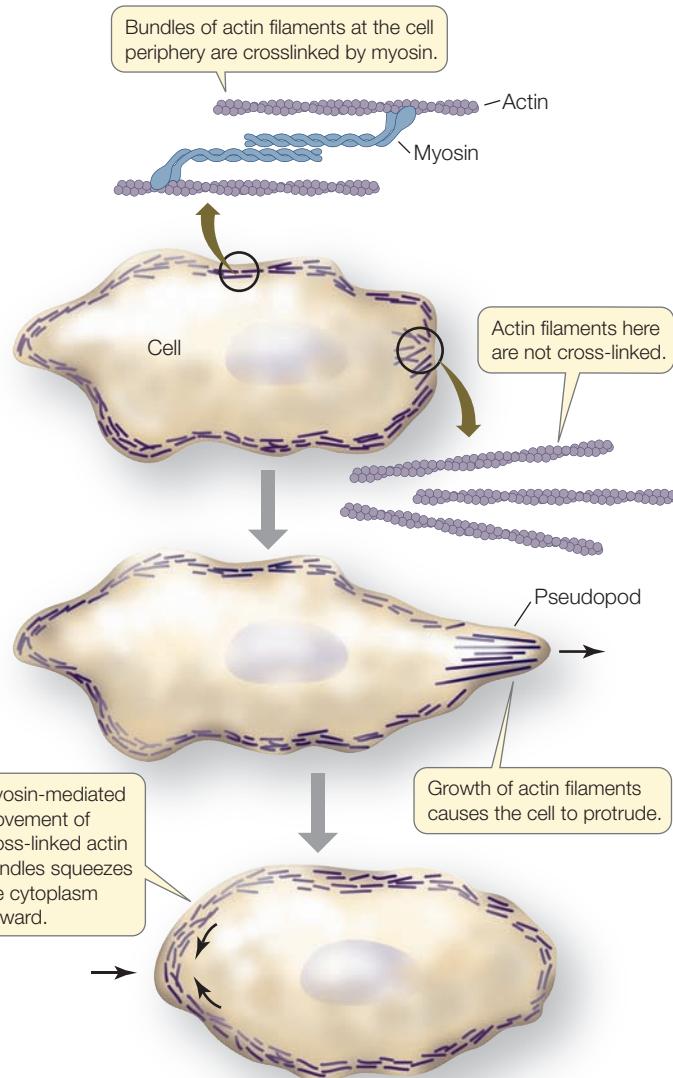
5.17 The Cytoskeleton Three highly visible and important structural components of the cytoskeleton are shown here in detail. These structures maintain and reinforce cell shape and contribute to cell movement.

microvilli that line the human intestine, giving it a larger surface area through which to absorb nutrients (**Figure 5.19**).

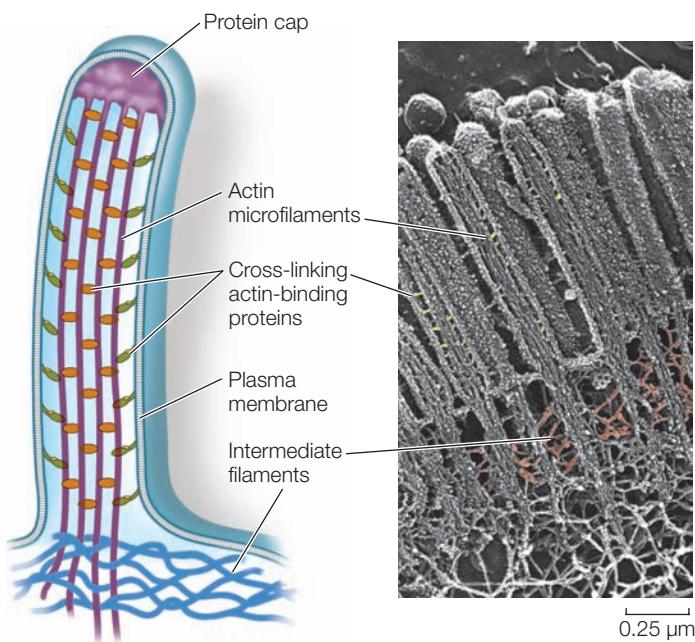
INTERMEDIATE FILAMENTS There are at least 50 different kinds of **intermediate filaments**, many of them specific to a few cell types. They generally fall into six molecular classes (based on amino acid sequence) that share the same general structure. One of these classes consists of fibrous proteins of the keratin family, which also includes the proteins that make up hair and fingernails. The intermediate filaments are tough, ropelike protein assemblages 8 to 12 nm in diameter (**Figure 5.17B**). Intermediate filaments are more permanent than the other two types; in cells they do not form and re-form, as the microtubules and microfilaments do.

Intermediate filaments have two major structural functions:

- They anchor cell structures in place. In some cells, intermediate filaments radiate from the nuclear envelope and help maintain the positions of the nucleus and other organelles in the cell. The lamins of the nuclear lamina are intermediate filaments (see Figure 5.8). Other kinds of intermediate filaments help hold in place the complex apparatus of microfilaments in the microvilli of intestinal cells (see Figure 5.19).
- They resist tension. For example, they maintain rigidity in body surface tissues by stretching through the cytoplasm and connecting specialized membrane structures called desmosomes (see Figure 6.7).



5.18 Microfilaments and Cell Movements Microfilaments mediate the movement of whole cells (as illustrated here for amoebic movement), as well as the movement of cytoplasm within a cell.



5.19 Microfilaments for Support Cells that line the intestine are folded into tiny projections called microvilli, which are supported by microfilaments. The microfilaments interact with intermediate filaments at the base of each microvillus. The microvilli increase the surface area of the cells, facilitating their absorption of small molecules.

MICROTUBULES The largest diameter components of the cytoskeletal system, **microtubules**, are long, hollow, unbranched cylinders about 25 nm in diameter and up to several micrometers long. Microtubules have two roles in the cell:

- They form a rigid internal skeleton for some cells.
- They act as a framework along which motor proteins can move structures within the cell.

Microtubules are assembled from dimers of the protein *tubulin*. A dimer is a molecule made up of two monomers. The polypeptide monomers that make up a tubulin dimer are known as α -tubulin and β -tubulin. Thirteen chains of tubulin dimers surround the central cavity of the microtubule (Figure 5.17C; see also Figure 5.20). The two ends of a microtubule are different: one is designated the plus (+) end, and the other the minus (-) end. Tubulin dimers can be rapidly added or subtracted, mainly at the plus end, lengthening or shortening the microtubule. This capacity to change length rapidly makes microtubules dynamic structures, permitting some animal cells to rapidly change shape.

Many microtubules radiate from a region of the cell called the microtubule organizing center. Tubulin polymerization results in a rigid structure, and tubulin depolymerization leads to its collapse.

In plants, microtubules help control the arrangement of the cellulose fibers of the cell wall. Electron micrographs of plants frequently show microtubules lying just inside the plasma membranes of cells that are forming or extending their cell walls. Experimental alteration of the orientation of these microtubules leads to a similar change in the cell wall and a new shape for the cell.

Microtubules serve as tracks for **motor proteins**, specialized molecules that use cellular energy to change their shape and move. Motor proteins bond to and move along the microtubules, carrying materials from one part of the cell to another. Microtubules are also essential in distributing chromosomes to daughter cells during cell division. Because of this, drugs such as vincristine and taxol that disrupt microtubule dynamics also disrupt cell division. These drugs are useful for treating cancer, where cell division is excessive.

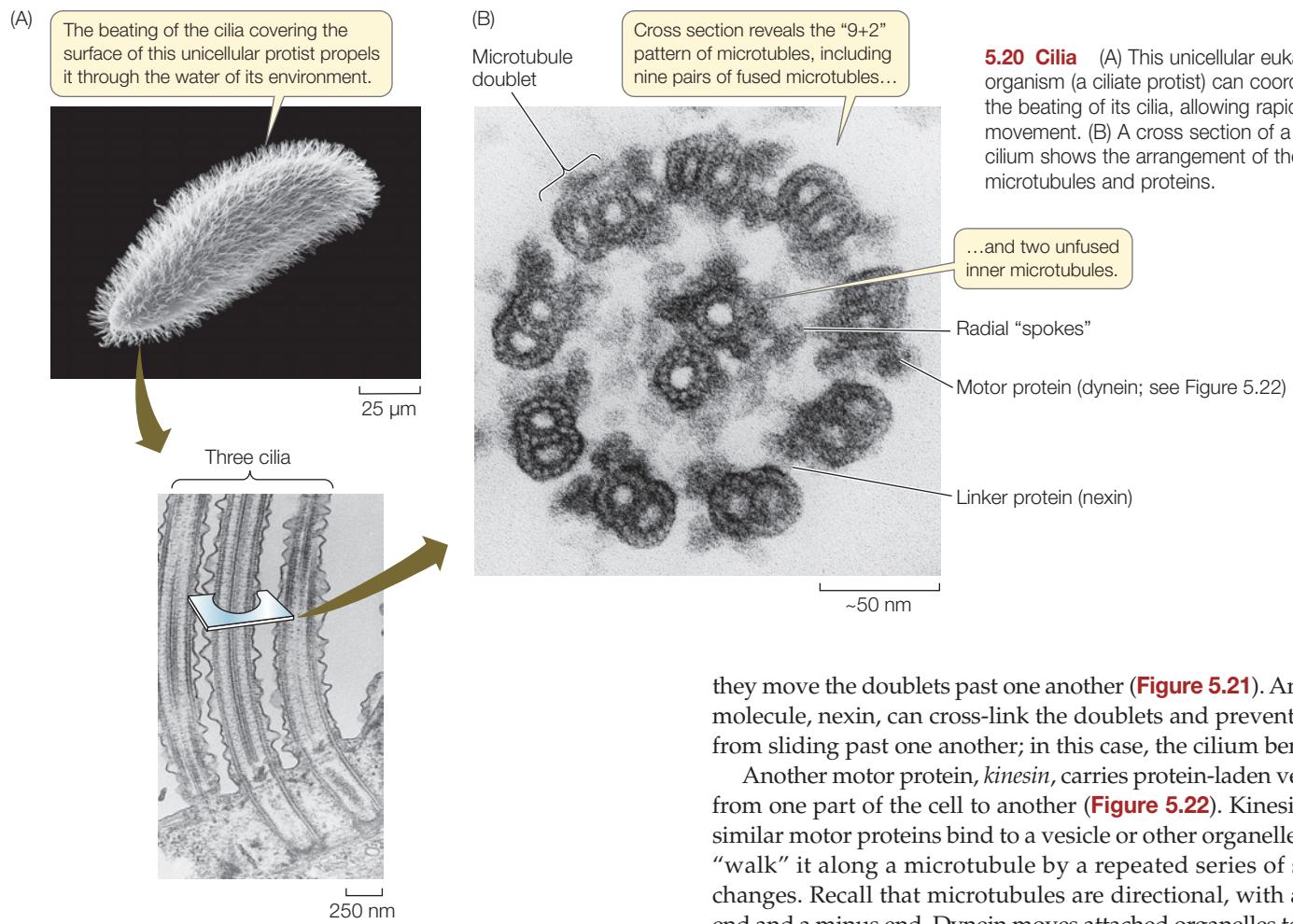
CILIA AND FLAGELLA Microtubules are also intimately associated with movable cell appendages: the **cilia** and **flagella**. Many eukaryotic cells have one or both of these appendages. Cilia are smaller than flagella—only 0.25 μm in length. They may move surrounding fluid over the surface of the cell (for example, protists or cells lining tubes through which eggs move, the oviducts). Eukaryotic flagella are 0.25 μm in diameter and 100–200 μm in length. (The structure and operation of eukaryotic flagella are very different from those of prokaryotic flagella; see Figure 5.5.) They may push or pull the cell through its aqueous environment (for example, protists or sperm). Cilia and eukaryotic flagella are both assembled from specialized microtubules and have identical internal structures, but differ in their length and pattern of beating:

- Cilia (singular *cilium*) are usually present in great numbers (Figure 5.20A). They beat stiffly in one direction and recover flexibly in the other direction (like a swimmer's arm), so that the recovery stroke does not undo the work of the power stroke.
- Eukaryotic flagella are usually found singly or in pairs. Waves of bending propagate from one end of a flagellum to the other in a snakelike undulation. Forces exerted by these waves on the surrounding fluid medium move the cell.

In cross section, a typical *cilium* or eukaryotic flagellum is surrounded by the plasma membrane and contains a “9 + 2” array of microtubules. As Figure 5.20B shows, nine fused pairs of microtubules—called doublets—form an outer cylinder, and one pair of unfused microtubules runs up the center. A spoke radiates from one microtubule of each doublet and connects the doublet to the center of the structure. These structures are essential to the bending motions of both cilia and flagella.

In the cytoplasm at the base of every eukaryotic flagellum and cilium is an organelle called a **basal body**. The nine microtubule doublets extend into the basal body. In the basal body, each doublet is accompanied by another microtubule, making nine sets of three microtubules. The central, unfused microtubules in the cilium do not extend into the basal body.

Centrioles are almost identical to the basal bodies of cilia and flagella. Centrioles are found in the microtubule organizing centers (sites of tubulin storage where microtubules polymerize) of all eukaryotes except the seed plants and some protists. Under the light microscope, a centriole looks like a small, featureless particle, but the electron microscope reveals that it contains a precise bundle of microtubules arranged in nine sets of three. Centrioles are involved in the formation of the mitotic spindle, to which the chromosomes attach during cell division (see Figure 11.10).



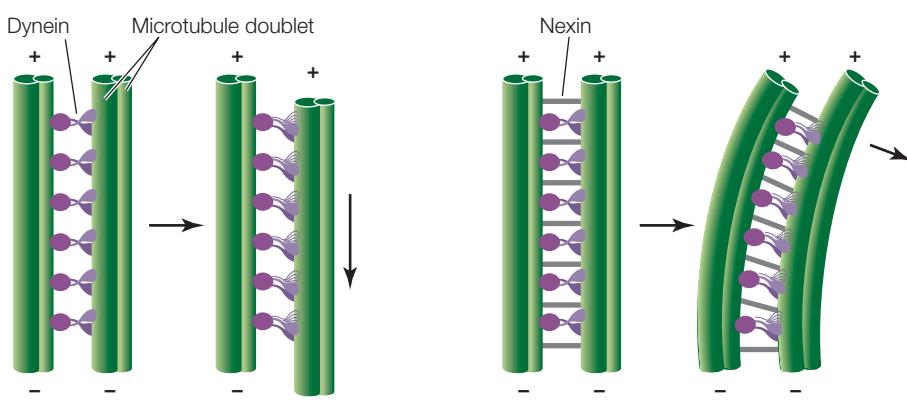
5.20 Cilia (A) This unicellular eukaryotic organism (a ciliate protist) can coordinate the beating of its cilia, allowing rapid movement. (B) A cross section of a single cilium shows the arrangement of the microtubules and proteins.

MOTOR PROTEINS AND MOVEMENT The nine microtubule doublets of cilia and flagella are linked by proteins. The motion of cilia and flagella results from the sliding of the microtubule doublets past each other. This sliding is driven by a motor protein called *dynein*, which can change its three-dimensional shape. All motor proteins work by undergoing reversible shape changes powered by energy from ATP hydrolysis. Dynein molecules that are attached to one microtubule doublet bind to a neighboring doublet. As the dynein molecules change shape, they move the doublets past one another (Figure 5.21). Another molecule, *nexin*, can cross-link the doublets and prevent them from sliding past one another; in this case, the cilium bends.

they move the doublets past one another (Figure 5.21). Another molecule, *nexin*, can cross-link the doublets and prevent them from sliding past one another; in this case, the cilium bends.

Another motor protein, *kinesin*, carries protein-laden vesicles from one part of the cell to another (Figure 5.22). Kinesin and similar motor proteins bind to a vesicle or other organelle, then “walk” it along a microtubule by a repeated series of shape changes. Recall that microtubules are directional, with a plus end and a minus end. Dynein moves attached organelles toward the minus end, while kinesin moves them toward the plus end (see Figure 5.17).

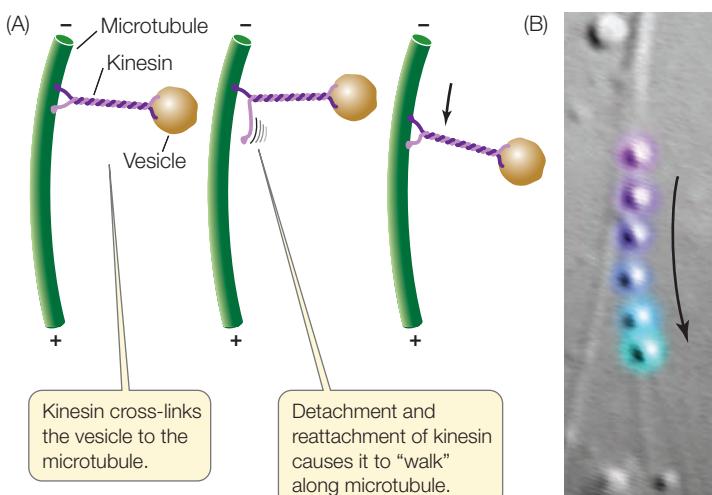
DEMONSTRATING CYTOSKELETON FUNCTIONS How do we know that the structural fibers of the cytoskeleton can achieve all these dynamic functions? We can observe an individual structure under the microscope and a function in a living cell that contains that structure. These observations may suggest that the structure carries out that function, but in science mere correlation does not show cause and effect. For example, light microscopy of living cells reveals that the cytoplasm is actively streaming around the cell, and that cytoplasm flows into an extended portion of an amoeboid cell during movement. The observed presence of cytoskeletal components suggests, but does not prove, their role in this process. Science seeks to show the specific links that relate one process, “A,” to a function, “B.” In cell biology, there are two ways to show that a structure or process “A” causes function “B”:



In isolated cilia without nexin cross-links, movement of dynein motor proteins causes microtubule doublets to slide past one another.

When nexin is present to cross-link the doublets, they cannot slide and the force generated by dynein movement causes the cilium to bend.

5.21 A Motor Protein Moves Microtubules in Cilia and Flagella A motor protein, dynein, causes microtubule doublets to slide past one another. In a flagellum or cilium, anchorage of the microtubule doublets to one another results in bending.



5.22 A Motor Protein Drives Vesicles along Microtubules

(A) Kinesin delivers vesicles or organelles to various parts of the cell by moving along microtubule “railroad tracks.” Kinesin moves things from the minus toward the plus end of a microtubule; dynein works similarly, but moves from the plus toward the minus end. (B) Powered by kinesin, a vesicle moves along a microtubule track in the protist *Dictyostelium*. The time sequence (time-lapse micrography at half-second intervals) is shown by the color gradient of purple to blue.

- **Inhibition:** use a drug that inhibits A and see if B still occurs. If it does not, then A is probably a causative factor for B. **Figure 5.23** shows an experiment with such a drug (an inhibitor) that demonstrates cause and effect in the case of the cytoskeleton and cell movement.
- **Mutation:** examine a cell that lacks the gene (or genes) for A and see if B still occurs. If it does not, then A is probably a causative factor for B. Part Four of this book describes many experiments using this genetic approach.

yourBioPortal.com

GO TO Animated Tutorial 5.2 • Eukaryotic Cell Tour

5.3 RECAP

The hallmark of eukaryotic cells is compartmentalization. Membrane-enclosed organelles process information, transform energy, form internal compartments for transporting proteins, and carry out intracellular digestion. An internal cytoskeleton plays several structural roles.

- What are some advantages of organelle compartmentalization? **See p. 84**
- Describe the structural and functional differences between rough and smooth endoplasmic reticulum. **See pp. 89–90 and Figure 5.10**
- Explain how motor proteins and microtubules move materials within the cell. **See pp. 95–98 and Figures 5.21 and 5.22**

INVESTIGATING LIFE

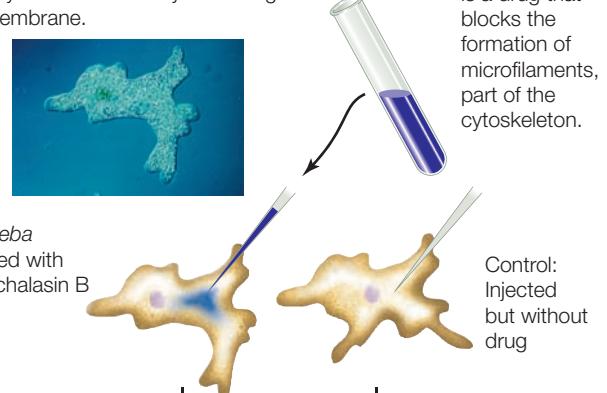
5.23 The Role of Microfilaments in Cell Movement—Showing Cause and Effect in Biology

After a test tube demonstration that the drug cytochalasin B prevented microfilament formation from monomeric precursors, the question was asked: Will the drug work like this in living cells and inhibit cell movement in *Amoeba*? Complementary experiments showed that the drug did not poison other cellular processes.

HYPOTHESIS Amoeboid cell movements are caused by the cytoskeleton.

METHOD

Amoeba proteus is a single-celled eukaryote that moves by extending its membrane.



RESULTS

Treated Amoeba rounds up and does not move



Control Amoeba continues to move



CONCLUSION Microfilaments of the cytoskeleton are essential for amoeboid cell movement.

FURTHER INVESTIGATION: The drug colchicine breaks apart microtubules. How would you show that these components of the cytoskeleton are not involved in cell movement in *Amoeba*?

Go to yourBioPortal.com for original citations, discussions, and relevant links for all INVESTIGATING LIFE figures.

All cells interact with their environments, and many eukaryotic cells are parts of multicellular organisms and must interact, and closely coordinate activities, with other cells. The plasma membrane plays a crucial role in these interactions, but other structures outside that membrane are involved as well.

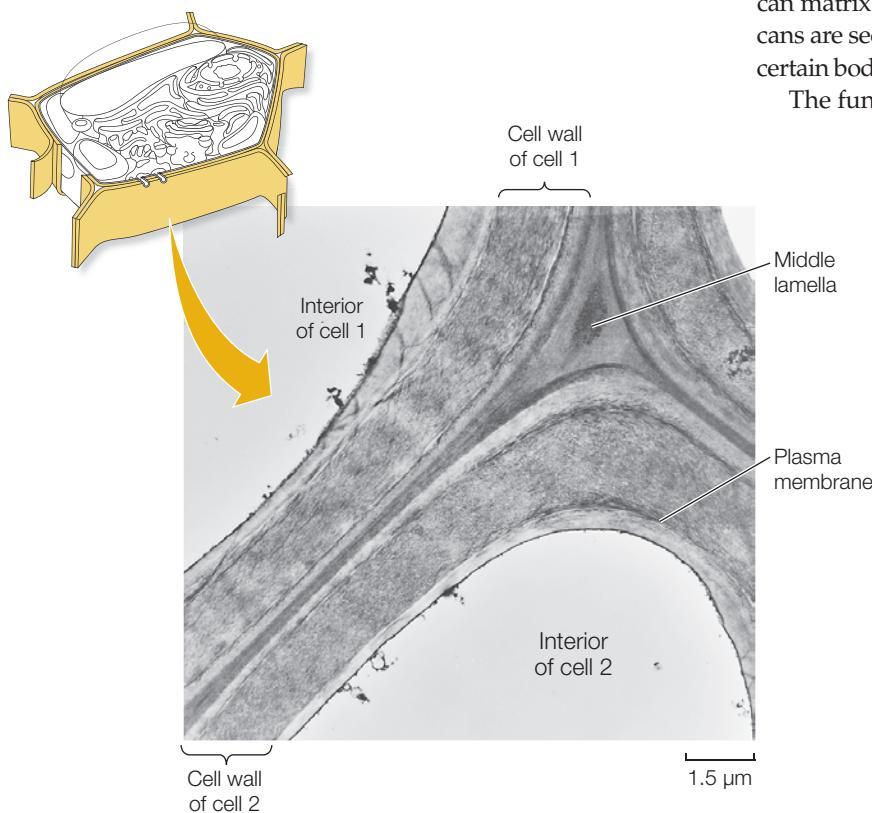
5.4 What Are the Roles of Extracellular Structures?

Although the plasma membrane is the functional barrier between the inside and the outside of a cell, many structures are produced by cells and secreted to the outside of the plasma membrane, where they play essential roles in protecting, supporting, or attaching cells to each other. Because they are outside the plasma membrane, these structures are said to be *extracellular*. The peptidoglycan cell wall of bacteria is an example of an extracellular structure (see Figure 5.4). In eukaryotes, other extracellular structures—the cell walls of plants and the extracellular matrices found between the cells of animals—play similar roles. Both of these structures are made up of two components:

- a prominent fibrous macromolecule
- a gel-like medium in which the fibers are embedded

The plant cell wall is an extracellular structure

The plant **cell wall** is a semirigid structure outside the plasma membrane (Figure 5.24). We consider the structure and role of the cell wall in more detail in Chapter 34. For now, we note that it is typical of a two-component extracellular matrix, with cellulose fibers (see Figure 3.16) embedded in other complex polysaccharides and proteins. The plant cell wall has three major roles:



5.24 The Plant Cell Wall The semirigid cell wall provides support for plant cells. It is composed of cellulose fibrils embedded in a matrix of polysaccharides and proteins.

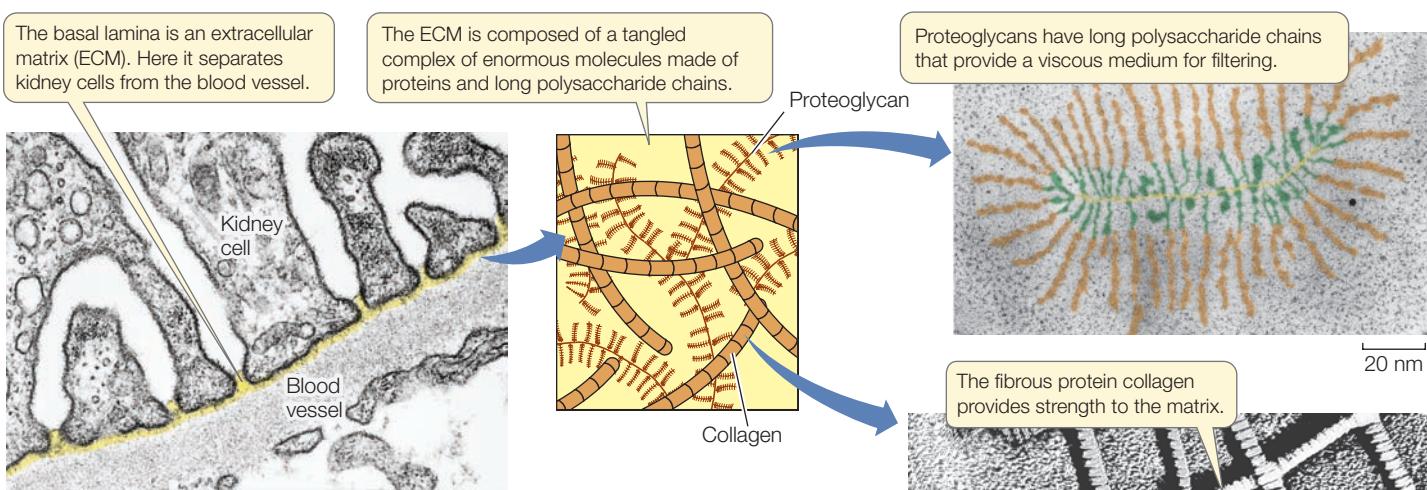
- It provides support for the cell and limits its volume by remaining rigid.
- It acts as a barrier to infection by fungi and other organisms that can cause plant diseases.
- It contributes to plant form by growing as plant cells expand. Because of their thick cell walls, plant cells viewed under a light microscope appear to be entirely isolated from one another. But electron microscopy reveals that this is not the case. The cytoplasms of adjacent plant cells are connected by numerous plasma membrane-lined channels, called **plasmodesmata**, that are about 20–40 nm in diameter and extend through the cell walls (see Figures 5.7 and 6.7). Plasmodesmata permit the diffusion of water, ions, small molecules, RNA, and proteins between connected cells, allowing for utilization of these substances far from their site of synthesis.

The extracellular matrix supports tissue functions in animals

Animal cells lack the semirigid wall that is characteristic of plant cells, but many animal cells are surrounded by, or in contact with, an **extracellular matrix**. This matrix is composed of three types of molecules: fibrous proteins such as **collagen** (the most abundant protein in mammals, constituting over 25 percent of the protein in the human body); a matrix of glycoproteins termed **proteoglycans**, consisting primarily of sugars; and a third group of proteins that link the fibrous proteins and the gel-like proteoglycan matrix together (Figure 5.25). These proteins and proteoglycans are secreted, along with other substances that are specific to certain body tissues, by cells that are present in or near the matrix.

The functions of the extracellular matrix are many:

- It holds cells together in tissues. In Chapter 6 we see how there is an intercellular “glue” that is involved in both cell recognition and adhesion.
- It contributes to the physical properties of cartilage, skin, and other tissues. For example, the mineral component of bone is laid down on an organized extracellular matrix.
- It helps filter materials passing between different tissues. This is especially important in the kidney.
- It helps orient cell movements during embryonic development and during tissue repair.
- It plays a role in chemical signaling from one cell to another. Proteins connect the cell’s plasma membrane to the extracellular matrix. These proteins (for example, *integrin*) span the plasma membrane and are involved with transmitting signals to the interior of the cell. This allows communication between the extracellular matrix and the cytoplasm of the cell.



5.25 An Extracellular Matrix Cells in the kidney secrete a basal lamina, an extracellular matrix that separates them from a nearby blood vessel and is also involved in filtering materials that pass between the kidney and the blood.

5.4 RECAP

Extracellular structures are produced by cells and secreted outside the plasma membrane. Most consist of a fibrous component in a gel-like medium.

- What are the functions of the cell wall in plants and the extracellular matrix in animals? See p. 100

We have now discussed the structures and some functions of prokaryotic and eukaryotic cells. Both exemplify the cell theory, showing that cells are the basic units of life and of biological continuity. Much of the rest of this part of the book will deal with these two aspects of cells. There is abundant evidence that the simpler prokaryotic cells are more ancient than eukaryotic cells, and that the first cells were probably prokaryotic. We now turn to the next step in cellular evolution, the origin of eukaryotic cells.

5.5 How Did Eukaryotic Cells Originate?

For about 2 billion years, life on Earth was entirely prokaryotic—from the time when prokaryotic cells first appeared until about 1.5 billion years ago, when eukaryotic cells arrived on the scene. The advent of compartmentalization—the hallmark of eukaryotes—was a major event in the history of life, as it permitted many more biochemical functions to coexist in the same cell than had previously been possible. Compared to the typical eukaryote, a single prokaryotic cell is often biochemically specialized, limited in the resources it can use and the functions it can perform.

What is the origin of compartmentalization? We will describe the evolution of eukaryotic organelles in more detail in Section 27.1. Here, we outline two major themes in this process.

Internal membranes and the nuclear envelope probably came from the plasma membrane

We noted earlier that some bacteria contain internal membranes. How could these arise? In electron micrographs, the internal

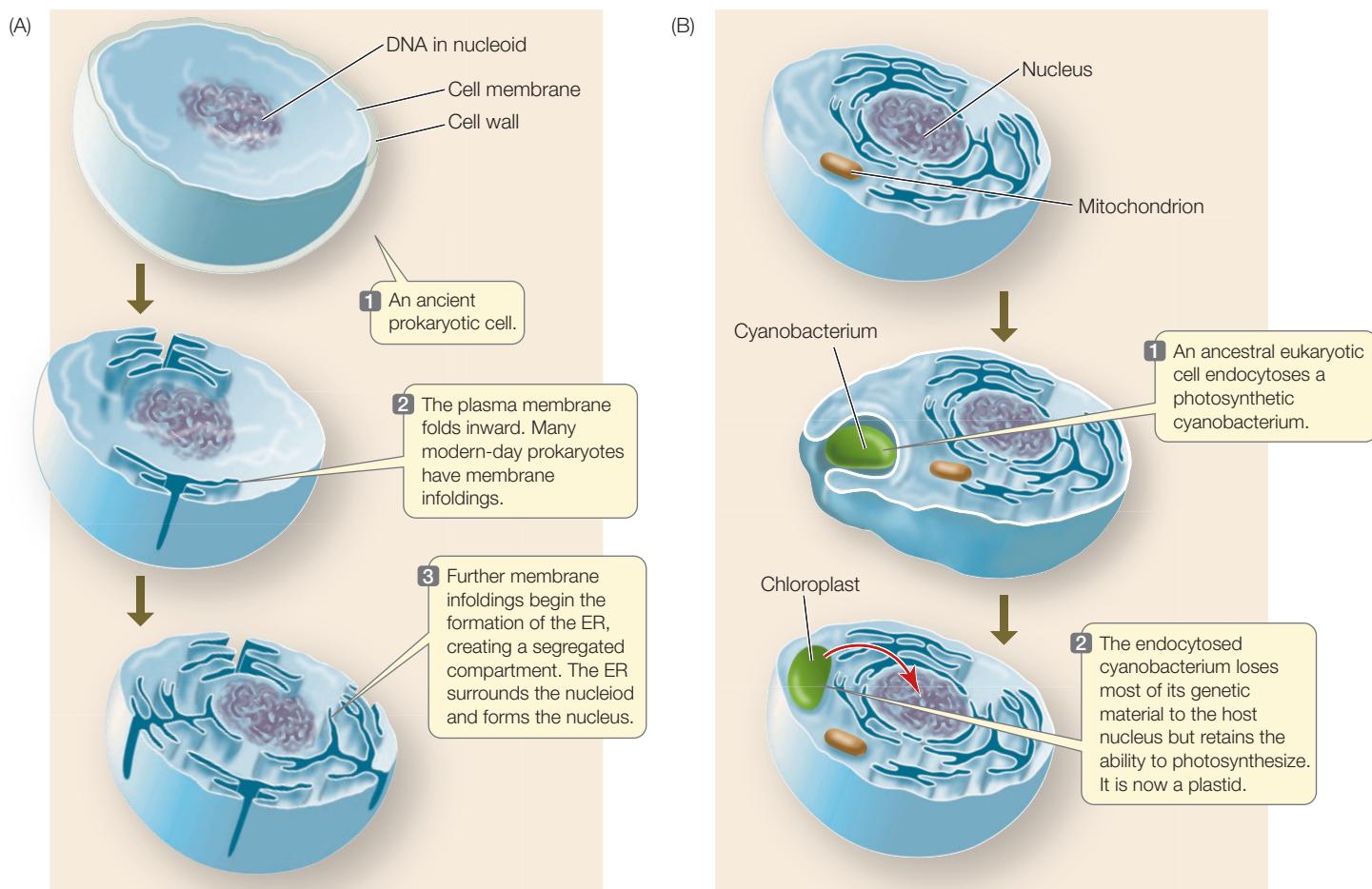
membranes of prokaryotes often appear to be inward folds of the plasma membrane. This has led to a theory that the endomembrane system and cell nucleus originated by a related process (**Figure 5.26A**). The close relationship between the ER and the nuclear envelope in today's eukaryotes is consistent with this theory.

A bacterium with enclosed compartments would have several evolutionary advantages. Chemicals could be concentrated within particular regions of the cell, allowing chemical reactions to proceed more efficiently. Biochemical activities could be segregated within organelles with, for example, a different pH from the rest of the cell, creating more favorable conditions for certain metabolic processes. Finally, gene transcription could be separated from translation, providing more opportunities for separate control of these steps in gene expression.

Some organelles arose by endosymbiosis

Symbiosis means “living together,” and often refers to two organisms that coexist, each one supplying something that the other needs. Biologists have proposed that some organelles—the mitochondria and the plastids—arose not by an infolding of the plasma membrane but by one cell ingesting another cell, giving rise to a symbiotic relationship. Eventually, the ingested cell lost its autonomy and some of its functions. In addition, many of the ingested cell’s genes were transferred to the host’s DNA. Mitochondria and plastids in today’s eukaryotic cells are the remnants of these *symbionts*, retaining some specialized functions that benefit their host cells. This is the essence of the **endosymbiosis theory** for the origin of organelles.

Consider the case of the plastid. About 2.5 billion years ago some prokaryotes (the cyanobacteria) developed photosynthesis (see Figure 1.9). The emergence of these prokaryotes was a key event in the evolution of complex organisms, because they increased the O₂ concentration in Earth’s atmosphere (see Section 1.2).



5.26 The Origin of Organelles (A) The endomembrane system and cell nucleus may have been formed by infolding and then fusion of the plasma membrane. (B) The endosymbiosis theory proposes that some organelles may be descended from prokaryotes that were engulfed by other, larger cells.

According to endosymbiosis theory, photosynthetic prokaryotes also provided the precursor of the modern-day plastid. Cells without cell walls can engulf relatively large particles by phagocytosis (see Figure 5.11). In some cases, such as that of phagocytes in the human immune system, the engulfed particle can be an entire cell, such as a bacterium. Plastids may have arisen by a similar event involving an ancestral eukaryote and a cyanobacterium (**Figure 5.26B**).

Among the abundant evidence supporting the endosymbiotic origin of plastids (see Section 27.1), perhaps the most remarkable comes from a sandy beach in Japan. Noriko Okamoto and Isao Inouye recently discovered a single-celled eukaryote that contains a large “chloroplast,” and named it *Hatena* (**Figure 5.27**). It turns out that the “chloroplast” is the remains of a green alga, *Nephroselmis*, which lives among the *Hatena* cells. When living autonomously, this algal cell has flagella, a cytoskeleton, ER, Golgi, and mitochondria in addition to a plastid. Once ingested by *Hatena*, all of these structures, and presumably their associated functions, are lost. What remains is essentially a plastid.

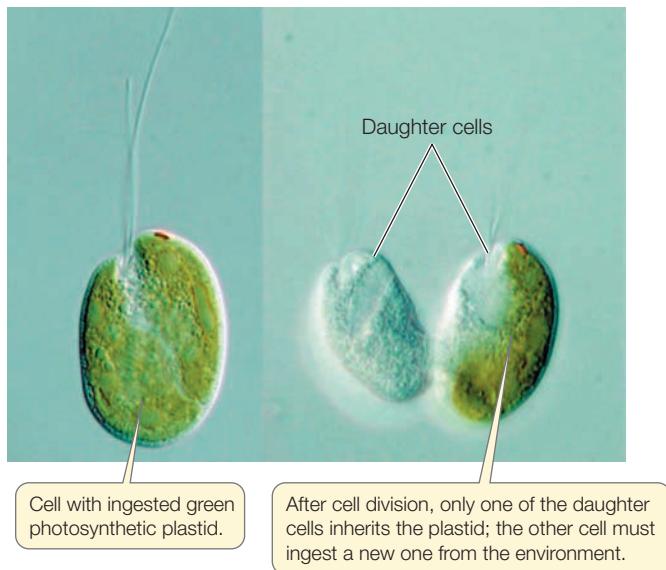
When *Hatena* divides, only one of the two daughter cells ends up with the “chloroplast.” The other cell finds and ingests its own *Nephroselmis* alga—almost like a “replay” of what may have occurred in the evolution of eukaryotic cells. No wonder the Japanese scientists call the host cell *Hatena*: in Japanese, it means “how odd”!

5.5 RECAP

Eukaryotic cells arose long after prokaryotic cells. Some organelles may have evolved by infolding of the plasma membrane, while others evolved by endosymbiosis.

- How could membrane infolding in a prokaryotic cell lead to the endomembrane system? **See p. 101 and Figure 5.26A**
- Explain the endosymbiosis theory for the origin of chloroplasts. **See Figure 5.26B**

In this chapter, we presented an overview of the structures of cells, with some ideas about their relationships and origins. As you now embark on the study of major cell functions, keep in



5.27 Endosymbiosis in Action A *Hatena* cell engulfs an algal cell, which then loses most of its cellular functions other than photosynthesis. This re-enacts a possible event in the origin of plastids in eukaryotic cells.

mind that the structures in a cell do not exist in isolation. They are part of a dynamic, interacting cellular system. In Chapter 6 we show that the plasma membrane is far from a passive barrier, but instead is a multi-functional system that connects the inside of the cell with its extracellular environment.

CHAPTER SUMMARY

5.1 What Features Make Cells the Fundamental Units of Life?

SEE WEB ACTIVITIES 5.1 AND 5.2

- All cells come from preexisting cells.
- Cells are small because a cell's surface area must be large compared with its volume to accommodate exchanges with its environment. **Review Figure 5.2**
- All cells are enclosed by a selectively permeable **plasma membrane** that separates their contents from the external environment.
- While certain biochemical processes, molecules, and structures are shared by all kinds of cells, two categories of cells—**prokaryotes** and **eukaryotes**—are easily distinguished.

5.2 What Features Characterize Prokaryotic Cells?

- Prokaryotic cells have no internal compartments, but have a **nucleoid** region containing DNA, and a **cytoplasm** containing **cytosol**, **ribosomes**, proteins, and small molecules. Some prokaryotes have additional protective structures, including a **cell wall**, an **outer membrane**, and a **capsule**. **Review Figure 5.4**
- Some prokaryotes have folded membranes that may be photosynthetic membranes, and some have **flagella** or **pili** for motility or attachment. **Review Figure 5.5**

5.3 What Features Characterize Eukaryotic Cells?

- Eukaryotic cells are larger than prokaryotic cells and contain many membrane-enclosed **organelles**. The membranes that envelop organelles ensure compartmentalization of their functions. **Review Figure 5.7**
- **Ribosomes** are sites of protein synthesis.
- The **nucleus** contains most of the cell's DNA and participates in the control of protein synthesis. **Review Figure 5.8**
- The **endomembrane system**—consisting of the **endoplasmic reticulum** and **Golgi apparatus**—is a series of interrelated compartments enclosed by membranes. It segregates proteins and

modifies them. **Lysosomes** contain many digestive enzymes. **Review Figures 5.10 and 5.11**, **WEB ACTIVITY 5.3**, **ANIMATED TUTORIAL 5.1**

- **Mitochondria** and **chloroplasts** are semi-autonomous organelles that process energy. Mitochondria are present in most eukaryotic organisms and contain the enzymes needed for cellular respiration. The cells of photosynthetic eukaryotes contain chloroplasts that harvest light energy for photosynthesis. **Review Figures 5.12 and 5.13**
- **Vacuoles** are prominent in many plant cells and consist of a membrane-enclosed compartment full of water and dissolved substances.
- The **microfilaments**, **intermediate filaments**, and **microtubules** of the **cytoskeleton** provide the cell with shape, strength, and movement. **Review Figure 5.18**

SEE ANIMATED TUTORIAL 5.2

5.4 What Are the Roles of Extracellular Structures?

- The plant **cell wall** consists principally of **cellulose**. Cell walls are pierced by **plasmodesmata** that join the cytoplasms of adjacent cells.
- In animals, the **extracellular matrix** consists of different kinds of proteins, including collagen and proteoglycans. **Review Figure 5.25**

5.5 How Did Eukaryotic Cells Originate?

- Infoldings of the plasma membrane could have led to the formation of some membrane-enclosed organelles, such as the endomembrane system and the nucleus. **Review Figure 5.26A**
- The **endosymbiosis theory** states that mitochondria and chloroplasts originated when larger prokaryotes engulfed, but did not digest, smaller prokaryotes. Mutual benefits permitted this symbiotic relationship to be maintained, allowing the smaller cells to evolve into the eukaryotic organelles observed today. **Review Figure 5.26B**

SELF-QUIZ

1. Which structure is generally present in both prokaryotic cells and eukaryotic plant cells?
 - a. Chloroplasts
 - b. Cell wall
 - c. Nucleus
 - d. Mitochondria
 - e. Microtubules
2. The major factor limiting cell size is the
 - a. concentration of water in the cytoplasm.
 - b. need for energy.
 - c. presence of membrane-enclosed organelles.
 - d. ratio of surface area to volume.
 - e. composition of the plasma membrane.
3. Which statement about mitochondria is *not* true?
 - a. The inner mitochondrial membrane folds to form cristae.
 - b. The outer membrane is relatively permeable to macromolecules.
 - c. Mitochondria are green because they contain chlorophyll.
 - d. Fuel molecules from the cytosol are used for respiration in mitochondria.
 - e. ATP is synthesized in mitochondria.
4. Which statement about plastids is true?
 - a. They are found in prokaryotes.
 - b. They are surrounded by a single membrane.
 - c. They are the sites of cellular respiration.
 - d. They are found only in fungi.
 - e. They may contain several types of pigments or polysaccharides.
5. If all the lysosomes within a cell suddenly ruptured, what would be the most likely result?
 - a. The macromolecules in the cytosol would break down.
 - b. More proteins would be made.
 - c. The DNA within mitochondria would break down.
 - d. The mitochondria and chloroplasts would divide.
 - e. There would be no change in cell function.
6. The Golgi apparatus
 - a. is found only in animals.
 - b. is found in prokaryotes.
 - c. is the appendage that moves a cell around in its environment.
 - d. is a site of rapid ATP production.
 - e. modifies and packages proteins.
7. Which structure is *not* surrounded by one or more membranes?
 - a. Ribosome
 - b. Chloroplast
 - c. Mitochondrion
 - d. Peroxisome
 - e. Vacuole
8. The cytoskeleton consists of
 - a. cilia, flagella, and microfilaments.
 - b. cilia, microtubules, and microfilaments.
 - c. internal cell walls.
 - d. microtubules, intermediate filaments, and microfilaments.
 - e. calcified microtubules.
9. Microfilaments
 - a. are composed of polysaccharides.
 - b. are composed of actin.
 - c. allow cilia and flagella to move.
 - d. make up the spindle that aids the movement of chromosomes.
 - e. maintain the position of the chloroplast in the cell.
10. Which statement about the plant cell wall is *not* true?
 - a. Its principal chemical components are polysaccharides.
 - b. It lies outside the plasma membrane.
 - c. It provides support for the cell.
 - d. It completely isolates adjacent cells from one another.
 - e. It is semirigid.

FOR DISCUSSION

1. The drug vincristine is used to treat many cancers. It apparently works by causing microtubules to depolymerize. Vincristine use has many side effects, including loss of dividing cells and nerve problems. Explain why this might be so.
2. Through how many membranes would a molecule have to pass in moving from the interior (stroma) of a chloroplast to the interior (matrix) of a mitochondrion? From the interior of a lysosome to the outside of a cell? From one ribosome to another?
3. How does the possession of double membranes by chloroplasts and mitochondria relate to the endosymbiosis theory of the origins of these organelles? What other evidence supports the theory?
4. Compare the extracellular matrix of the animal cell with the plant cell wall, with respect to composition of the fibrous and nonfibrous components, rigidity, and connectivity of cells.

ADDITIONAL INVESTIGATION

The pathway of newly synthesized proteins can be followed through the cell using a “pulse-chase” experiment. During synthesis, proteins are tagged with a radioactive isotope (the “pulse”), and then the cell is allowed to process the proteins for varying periods of time. The locations of the radioactive pro-

teins are then determined by isolating cell organelles and quantifying their radioactivity. How would you use this method, and what results would you expect for (a) a lysosomal enzyme and (b) a protein that is released from the cell?

6

Cell Membranes

Membranes and memory

James noticed the changes in his grandfather when he was home from college for the winter holiday. He and grandpa John had always joked about grandpa John's missing keys and glasses; the old man, who had lived with James' family since his wife died, was forever searching for them. Now the memory lapses had become more pronounced. When James introduced his new girlfriend to the family, he was relieved (as was she) when she was welcomed with open arms. But an hour later, grandpa John just stared at her, unable to remember who she was. By the time James came home for the summer, his grandfather had become withdrawn; he could no longer talk about current events, and often he became confused and lashed out in anger.

James' grandfather had Alzheimer's disease. This condition is most common in (but not limited to) the elderly, and as more people today are living to advanced ages, more and more Alzheimer's cases are diagnosed. But the symptoms are not new to human experience or to medicine. The condition was first recognized as a disease in 1901. That year, the family of 51-year-old Frau Auguste D. brought her to Dr. Alois Alzheimer at the Frankfurt hospital in Germany. She had severe memory lapses, accused her husband of infidelity, and had difficulty communicating. These symptoms got worse before she died several years later. When Alzheimer autopsied her brain, he saw that the parts of the brain that are important in thought and speech were shrunken. Moreover, when he examined these areas

through the microscope he saw abnormal protein deposits in and around the brain cells.

In the century since Alzheimer's original case, cell biologists have investigated the nature of these abnormal deposits, now known as *plaques*. It turns out that the key events that produce plaques take place in the plasma membrane of nerve cells in the brain. Plaques are clumps of the protein amyloid beta, which at high levels is toxic to brain cells. Amyloid beta is a small piece of a larger amyloid precursor protein (APP), which is embedded in the nerve cell plasma membrane; APP is cut twice by two other membrane proteins, β -secretase and γ -secretase, to produce amyloid beta, which is released from the membrane to fall outside of the cell. All these proteins are present in a variety of animal cells and have multiple important



Dr. Alzheimer's Patient Frau Auguste D., who died in 1906, was the first patient described with progressive dementia by Dr. Alois Alzheimer.



Plaques in the Brain At autopsy, the brain of an Alzheimer's disease patient accumulates plaques (dark fibers in this micrograph) composed of protein fragments produced by an enzyme in the nerve cell membrane.

roles in the dynamic cell membrane; they may even be essential for normal nervous system development and function.

So what goes wrong in Alzheimer's disease? Cells in the diseased brain might be producing too much amyloid beta (e.g., because γ -secretase is too active) or producing it at the wrong time (e.g., in old age instead of infancy). One form of the disease is caused by a mutant form of γ -secretase, which has a tendency to cut APP in the "wrong" place, thereby producing a particularly toxic form of amyloid beta. Because of their role in producing plaques, APP and γ -secretase are potential targets for Alzheimer's disease therapies.

Learning how membranes are made and how they work has been a key to understanding, and perhaps treating, this increasingly prevalent disease.

IN THIS CHAPTER we focus on the structure and functions of biological membranes. First we describe the composition and structure of biological membranes. We go on to discuss their functions—how membranes are involved in intercellular interactions, and how membranes regulate which substances enter and leave the cell.

CHAPTER OUTLINE

- 6.1** What Is the Structure of a Biological Membrane?
- 6.2** How Is the Plasma Membrane Involved in Cell Adhesion and Recognition?
- 6.3** What Are the Passive Processes of Membrane Transport?
- 6.4** What are the Active Processes of Membrane Transport?
- 6.5** How Do Large Molecules Enter and Leave a Cell?
- 6.6** What Are Some Other Functions of Membranes?

6.1 What Is the Structure of a Biological Membrane?

The physical organization and functioning of all biological membranes depend on their constituents: lipids, proteins, and carbohydrates. You are already familiar with these molecules from Chapter 3; it may be useful to review that chapter now. The lipids establish the physical integrity of the membrane and create an effective barrier to the rapid passage of hydrophilic materials such as water and ions. In addition, the phospholipid bilayer serves as a lipid "lake" in which a variety of proteins "float" (Figure 6.1). This general design is known as the **fluid mosaic model**.

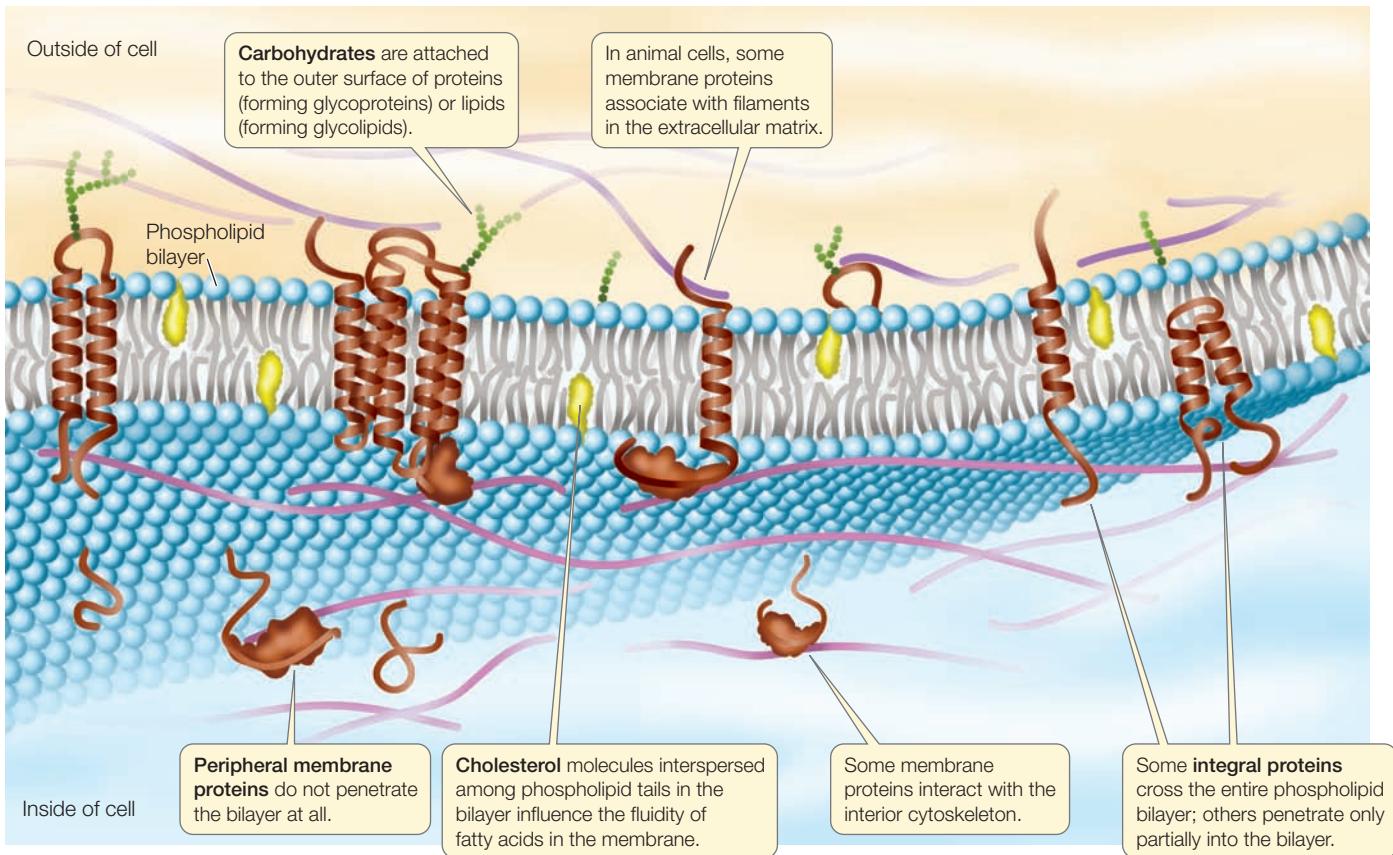
In the fluid mosaic model for biological membranes, the proteins are noncovalently embedded in the phospholipid bilayer by their hydrophobic regions (or *domains*), but their hydrophilic domains are exposed to the watery conditions on either side of the bilayer. These membrane proteins have a number of functions, including moving materials through the membrane and receiving chemical signals from the cell's external environment. Each membrane has a set of proteins suitable for the specialized functions of the cell or organelle it surrounds.

The carbohydrates associated with membranes are attached either to the lipids or to protein molecules. In plasma membranes, carbohydrates are located on the outside of the cell, where they may interact with substances in the external environment. Like some of the membrane proteins, carbohydrates are crucial in recognizing specific molecules, such as those on the surfaces of adjacent cells.

Although the fluid mosaic model is largely valid for membrane structure, it does not say much about membrane composition. As you read about the different molecules in membranes in the next sections, keep in mind that some membranes have more protein than lipids, others are lipid-rich, others have significant amounts of cholesterol or other sterols, and still others are rich in carbohydrates.

Lipids form the hydrophobic core of the membrane

The lipids in biological membranes are usually *phospholipids*. Recall from Section 2.2 that some compounds are hydrophilic ("water-loving") and others are hydrophobic ("water-hating"), and from Section 3.4 that a phospholipid molecule has regions of both kinds:



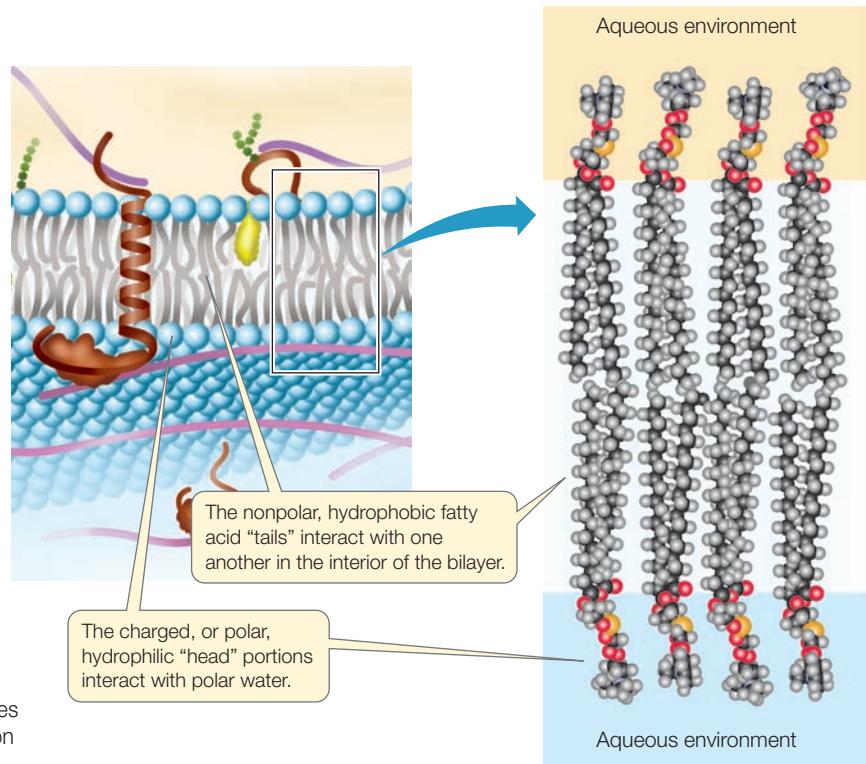
6.1 The Fluid Mosaic Model The general molecular structure of biological membranes is a continuous phospholipid bilayer which has proteins embedded in or associated with it.

yourBioPortal.com
GO TO Web Activity 6.1 • The Fluid-Mosaic Model

- **Hydrophilic regions:** The phosphorus-containing “head” of the phospholipid is electrically charged and therefore associates with polar water molecules.
- **Hydrophobic regions:** The long, nonpolar fatty acid “tails” of the phospholipid associate with other nonpolar materials, but they do not dissolve in water or associate with hydrophilic substances.

Because of these properties, one way in which phospholipids can coexist with water is to form a **bilayer**, with the fatty acid “tails” of the two layers interacting with each other and the polar “heads” facing the outside aqueous environment (**Figure 6.2**). The thickness of a biological membrane is about 8 nm (0.008 μm), which is twice the length of a typical phospholipid—another indication that the membrane consists of a lipid bilayer. This thickness is about 8,000 times thinner than a piece of paper.

6.2 A Phospholipid Bilayer The phospholipid bilayer separates two aqueous regions. The eight phospholipid molecules shown on the right represent a small cross section of a membrane bilayer.



In the laboratory, it is easy to make artificial bilayers with the same organization as natural membranes. Small holes in such bilayers seal themselves spontaneously. This capacity of lipids to associate with one another and maintain a bilayer organization helps biological membranes to fuse during vesicle formation, phagocytosis, and related processes.

All biological membranes have a similar structure, but differ in the kinds of proteins and lipids they contain. Membranes from different cells or organelles may differ greatly in their *lipid composition*. Not only are phospholipids highly variable, but a significant proportion of the lipid content in an animal cell membrane may be cholesterol.

Phospholipids can differ in terms of fatty acid chain length (number of carbon atoms), degree of unsaturation (double bonds) in the fatty acids, and the polar (phosphate-containing) groups present. The most common fatty acids with their chain length and degree of unsaturation are:

- Palmitic: C₁₄, no double bonds, saturated
- Palmitoleic: C₁₆, one double bond
- Stearic: C₁₈, no double bonds, saturated
- Oleic: C₁₈, one double bond
- Linoleic: C₁₈, two double bonds
- Linolenic: C₁₈, three double bonds

The saturated fatty acid chains allow close packing of fatty acids in the bilayer, while the “kinks” in unsaturated fatty acids (see Figure 3.19) make for a less dense, more fluid packing. These less-dense membranes in animal cells can accommodate cholesterol molecules.

Up to 25 percent of the lipid content of an animal cell plasma membrane may be cholesterol. When present, cholesterol is important for membrane integrity; the cholesterol in your membranes is not hazardous to your health. A molecule of cholesterol is usually situated next to an unsaturated fatty acid.

The phospholipid bilayer stabilizes the entire membrane structure, but leaves it flexible. The fatty acids of the phospholipids make the hydrophobic interior of the membrane somewhat fluid—about as fluid as lightweight machine oil. This fluidity permits some molecules to move laterally within the plane of the membrane. A given phospholipid molecule in the plasma membrane can travel from one end of the cell to the other in a little more than a second! On the other hand, seldom does a phospholipid molecule in one half of the bilayer spontaneously flip over to the other side. For that to happen, the polar part of the molecule would have to move through the hydrophobic interior of the membrane. Since spontaneous phospholipid flip-flops are rare, the inner and outer halves of the bilayer may be quite different in the kinds of phospholipids they contain.

The fluidity of a membrane is affected by its lipid composition and by its temperature. Long-chain, saturated fatty acids pack tightly beside one another, with little room for movement. Cholesterol interacts hydrophobically with the fatty acid chains. A membrane with these components is less fluid than one with shorter-chain fatty acids, unsaturated fatty acids, or less cholesterol. Adequate membrane fluidity is essential for many of the

functions we will describe in this chapter. Because molecules move more slowly and fluidity decreases at reduced temperatures, membrane functions may decline under cold conditions in organisms that cannot keep their bodies warm. To address this problem, some organisms simply change the lipid composition of their membranes when they get cold, replacing saturated with unsaturated fatty acids and using fatty acids with shorter tails. These changes play a role in the survival of plants, bacteria, and hibernating animals during the winter.

Membrane proteins are asymmetrically distributed

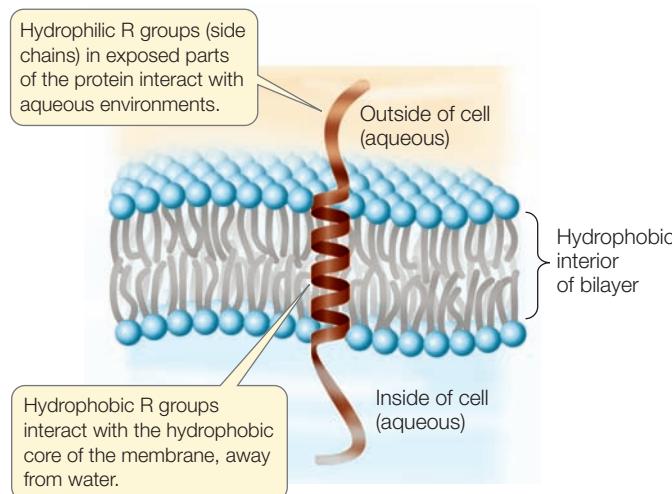
All biological membranes contain proteins. Typically, plasma membranes have one protein molecule for every 25 phospholipid molecules. This ratio varies depending on membrane function. In the inner membrane of the mitochondrion, which is specialized for energy processing, there is one protein for every 15 lipids. On the other hand, myelin—a membrane that encloses portions of some neurons (nerve cells) and acts as an electrical insulator—has only one protein for every 70 lipids.

There are two general types of membrane proteins: peripheral proteins and integral proteins.

Peripheral membrane proteins lack exposed hydrophobic groups and are not embedded in the bilayer. Instead, they have polar or charged regions that interact with exposed parts of integral membrane proteins, or with the polar heads of phospholipid molecules (see Figure 6.1).

Integral membrane proteins are at least partly embedded in the phospholipid bilayer (see Figure 6.1). Like phospholipids, these proteins have both hydrophilic and hydrophobic regions (Figure 6.3).

- *Hydrophilic domains*: Stretches of amino acids with hydrophilic side chains (see Table 3.1) give certain regions of the



6.3 Interactions of Integral Membrane Proteins An integral membrane protein is held in the membrane by the distribution of the hydrophilic and hydrophobic side chains on its amino acids. The hydrophilic parts of the protein extend into the aqueous cell exterior and the internal cytoplasm. The hydrophobic side chains interact with the hydrophobic lipid core of the membrane.

protein a polar character. These hydrophilic domains interact with water and stick out into the aqueous environment inside or outside the cell.

- **Hydrophobic domains:** Stretches of amino acids with hydrophobic side chains give other regions of the protein a nonpolar character. These domains interact with the fatty acids in the interior of the phospholipid bilayer, away from water.

A special preparation method for electron microscopy, called **freeze-fracturing**, reveals proteins that are embedded in the phospholipid bilayers of cellular membranes (**Figure 6.4**). When the two lipid *leaflets* (or layers) that make up the bilayer are separated, the proteins can be seen as bumps that protrude from the interior of each membrane. The bumps are not observed when artificial bilayers of pure lipid are freeze-fractured.

According to the fluid mosaic model, the proteins and lipids in a membrane are somewhat independent of each other and interact only noncovalently. The polar ends of proteins can interact with the polar ends of lipids, and the nonpolar regions of both molecules can interact hydrophobically.

However, some membrane proteins have fatty acids or other lipid groups covalently attached to them. Proteins in this subgroup of integral membrane proteins are referred to as *anchored membrane proteins*, because their hydrophobic lipid components allow them to insert themselves into the phospholipid bilayer.

Proteins are asymmetrically distributed on the inner and outer surfaces of membranes. An integral protein that extends all the way through the phospholipid bilayer and protrudes on both sides is known as a **transmembrane protein**. In addition to one or more *transmembrane domains* that extend through the bilayer, such a protein may have domains with other specific functions on the inner and outer sides of the membrane. Peripheral membrane proteins are localized on one side of the membrane or the other. This asymmetrical arrangement of membrane proteins gives the two surfaces of the membrane different properties. As we will soon see, these differences have great functional significance.

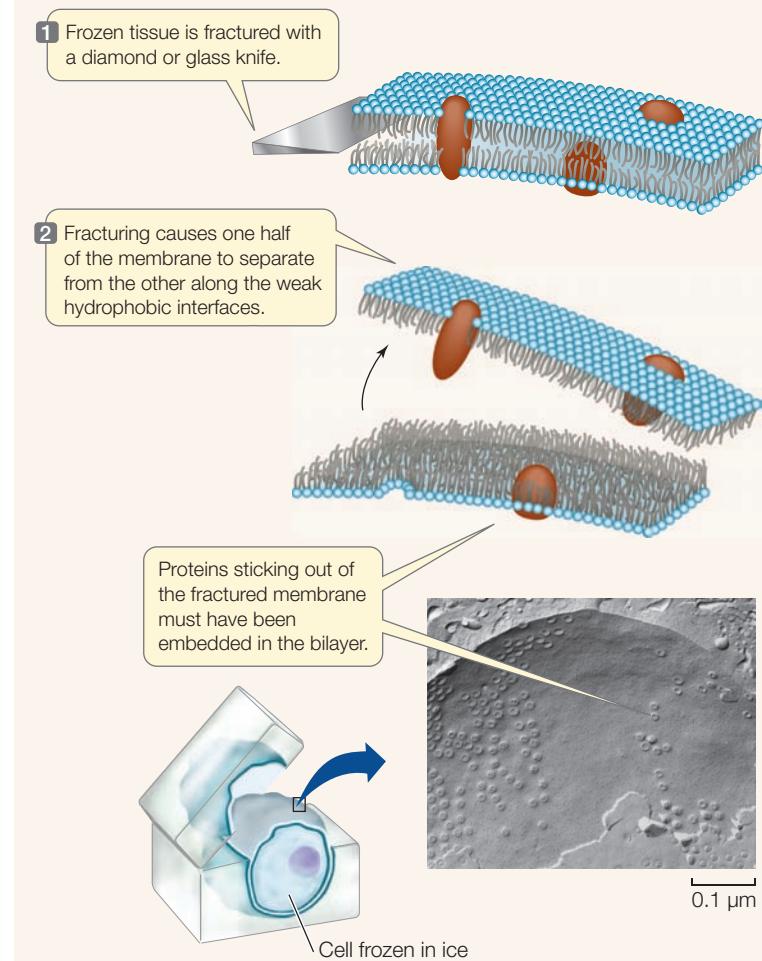
Like lipids, some membrane proteins move around relatively freely within the phospholipid bilayer. Experiments that involve the technique of cell fusion illustrate this migration dramatically. When two cells are fused, a single continuous membrane forms and surrounds both cells, and some proteins from each cell distribute themselves uniformly around this membrane (**Figure 6.5**).

Although some proteins are free to migrate in the membrane, others are not, but rather appear to be “anchored” to a specific region of the membrane. These membrane regions are like a corral of horses on a farm: the horses are free to move around within the fenced area, but not outside it. An example is the protein in the plasma membrane of a muscle cell that recognizes a chemical signal from a neuron. This protein is normally found only at the specific region where the neuron meets the muscle cell.

TOOLS FOR INVESTIGATING LIFE

6.4 Membrane Proteins Revealed by the Freeze-Fracture Technique

This HeLa cell (a human cell) membrane was first frozen to immobilize the lipids and proteins, and then fractured so that the bilayer was split open.



Proteins inside the cell can restrict the movement of proteins within a membrane. The cytoskeleton may have components just below the inner face of the membrane that are attached to membrane proteins protruding into the cytoplasm. The stability of the cytoskeletal components may thus restrict movement of attached membrane proteins.

Membranes are constantly changing

Membranes in eukaryotic cells are constantly forming, transforming from one type to another, fusing with one another, and breaking down. As we discuss in Chapter 5, fragments of membrane move, in the form of vesicles, from the endoplasmic reticulum (ER) to the Golgi, and from the Golgi to the plasma membrane (see Figure 5.10). Secondary lysosomes form when primary lysosomes from the Golgi fuse with phagosomes from the plasma membrane (see Figure 5.11).

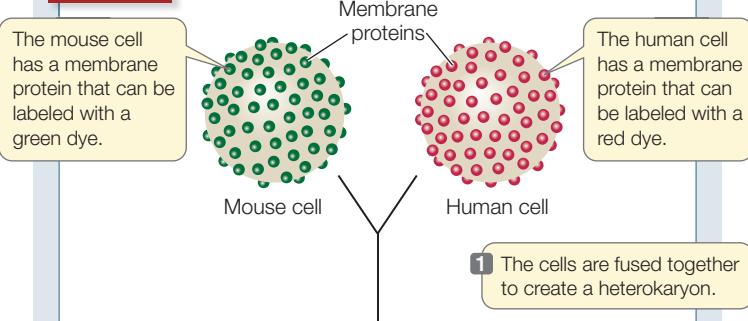
INVESTIGATING LIFE

6.5 Rapid Diffusion of Membrane Proteins

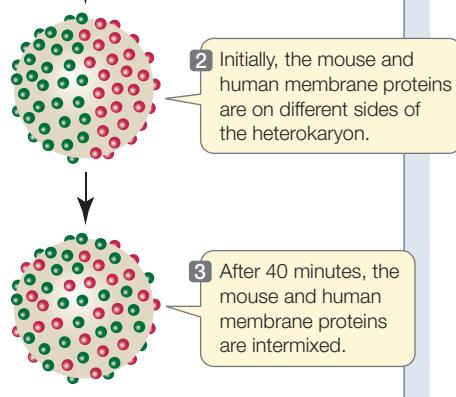
Two animal cells can be fused together in the laboratory, forming a single large cell (heterokaryon). This phenomenon was used to test whether membrane proteins can diffuse independently in the plane of the plasma membrane.

HYPOTHESIS Proteins embedded in a membrane can diffuse freely within the membrane.

METHOD



RESULTS



CONCLUSION Membrane proteins can diffuse rapidly in the plane of the membrane.

Go to [yourBioPortal.com](#) for original citations, discussions, and relevant links for all INVESTIGATING LIFE figures.

Because all membranes appear similar under the electron microscope, and because they interconvert readily, we might expect all subcellular membranes to be chemically identical. However, that is not the case, for there are major chemical differences among the membranes of even a single cell. Membranes are changed chemically when they form parts of certain organelles. In the Golgi apparatus, for example, the membranes of the *cis* face closely resemble those of the endoplasmic reticulum in chemical composition, but those of the *trans* face are more similar to the plasma membrane.

Plasma membrane carbohydrates are recognition sites

In addition to lipids and proteins, the plasma membrane contains carbohydrates. The carbohydrates are located on the outer

surface of the plasma membrane and serve as recognition sites for other cells and molecules, as you will see in Section 6.2.

Membrane-associated carbohydrates may be covalently bonded to lipids or to proteins:

- A **glycolipid** consists of a carbohydrate covalently bonded to a lipid. Extending outside the cell surface, the carbohydrate may serve as a recognition signal for interactions between cells. For example, the carbohydrates on some glycolipids change when cells become cancerous. This change may allow white blood cells to target cancer cells for destruction.
- A **glycoprotein** consists of a carbohydrate covalently bonded to a protein. The bound carbohydrate is an oligosaccharide, usually not exceeding 15 monosaccharide units in length (see Section 3.3). The oligosaccharides of glycoproteins often function as signaling sites, as do the carbohydrates attached to glycolipids.

The “alphabet” of monosaccharides on the outer surfaces of membranes can generate a large diversity of messages. Recall from Section 3.3 that sugar molecules consist of three to seven carbons that are attached at different sites to one another. They may form linear or branched oligosaccharides with many different three-dimensional shapes. An oligosaccharide of a specific shape on one cell can bind to a complementary shape on an adjacent cell. This binding is the basis of cell–cell adhesion.

6.1 RECAP

The fluid mosaic model applies to both the plasma membrane and the membranes of organelles. An integral membrane protein has both hydrophilic and hydrophobic domains, which affect its position and function in the membrane. Carbohydrates that attach to lipids and proteins on the outside of the membrane serve as recognition sites.

- What are some of the features of the fluid mosaic model of biological membranes?
See p. 106
- Explain how the hydrophobic and hydrophilic regions of phospholipids cause a membrane bilayer to form. **See Figures 6.1 and 6.2**
- What differentiates an integral protein from a peripheral protein? **See p. 108 and Figure 6.1**
- What is the experimental evidence that membrane proteins can diffuse in the plane of the membrane?
See pp. 109–110 and Figure 6.5

Now that you understand the structure of biological membranes, let’s see how their components function. In the next section we’ll focus on the membrane that surrounds individual cells: the plasma membrane. We’ll look at how the plasma membrane allows individual cells to be grouped together into multicellular systems of tissues.

6.2 How Is the Plasma Membrane Involved in Cell Adhesion and Recognition?

Some organisms, such as bacteria, are unicellular; that is, the entire organism is a single cell. Others, such as plants and animals, are multicellular—composed of many cells. Often these cells exist in specialized groups with similar functions, called tissues. Your body has about 60 trillion cells, arranged in different kinds of tissues (such as muscle, nerve, and epithelium).

Two processes allow cells to arrange themselves in groups:

- **Cell recognition**, in which one cell specifically binds to another cell of a certain type
- **Cell adhesion**, in which the connection between the two cells is strengthened

Both processes involve the plasma membrane. They are most easily studied if a tissue is separated into its individual cells, which are then allowed to adhere to one another again. Simple organisms provide a good model for studying processes that also occur in the complex tissues of larger species. Studies of sponges, for example, have revealed how cells associate with one another.

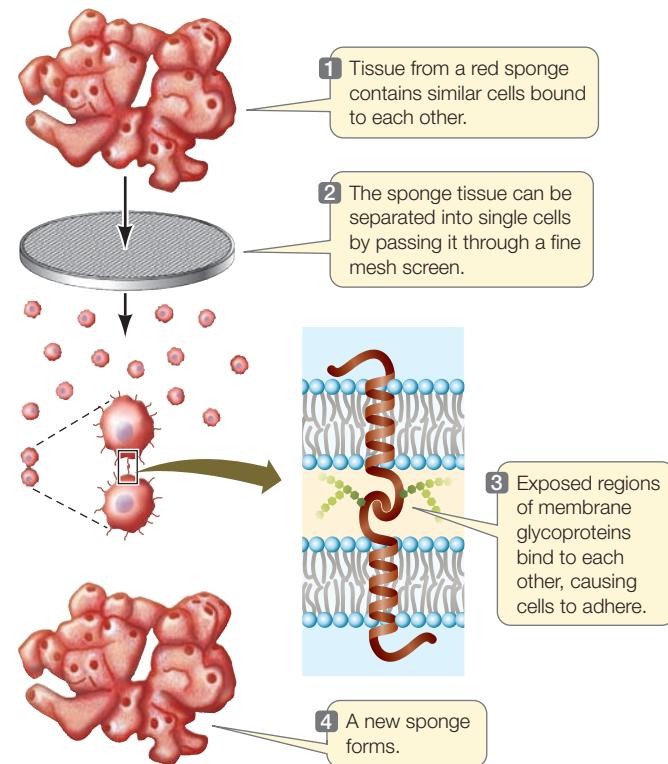
A sponge is a multicellular marine animal with a simple body plan that consists of only a few distinct tissues (see Section 31.5). The cells of a sponge adhere to one another, but can be separated mechanically by passing the animal several times through a fine wire screen (**Figure 6.6**). Through this process, what was a single animal becomes hundreds of individual cells suspended in seawater. Remarkably, if the cell suspension is shaken for a few hours, the cells bump into one another and stick together in the same shape and organization as the original sponge! The cells recognize and adhere to one another, and re-form the original tissues.

There are many different species of sponges. If disaggregated sponge cells from two different species are placed in the same container and shaken, individual cells will stick only to other cells of the same species. Two different sponges form, just like the ones at the start of the experiment. This demonstrates not just adhesion, but species-specific cell recognition.

Such tissue-specific and species-specific cell recognition and cell adhesion are essential to the formation and maintenance of tissues and multicellular organisms. Think of your own body. What keeps muscle cells bound to muscle cells and skin to skin? Specific cell adhesion is so obvious a characteristic of complex organisms that it is easy to overlook. You will see many examples of specific cell adhesion throughout this book; here, we describe its general principles. As you will see, cell recognition and cell adhesion depend on plasma membrane proteins.

Cell recognition and cell adhesion involve proteins at the cell surface

The molecule responsible for cell recognition and adhesion in sponges is a huge integral membrane glycoprotein (which is 80 percent carbohydrate by molecular weight) that is partly embedded in the plasma membrane, with the carbohydrate part sticking out and exposed to the environment (and to other



6.6 Cell Recognition and Adhesion In most cases (including the aggregation of animal cells into tissues), protein binding is homotypic.

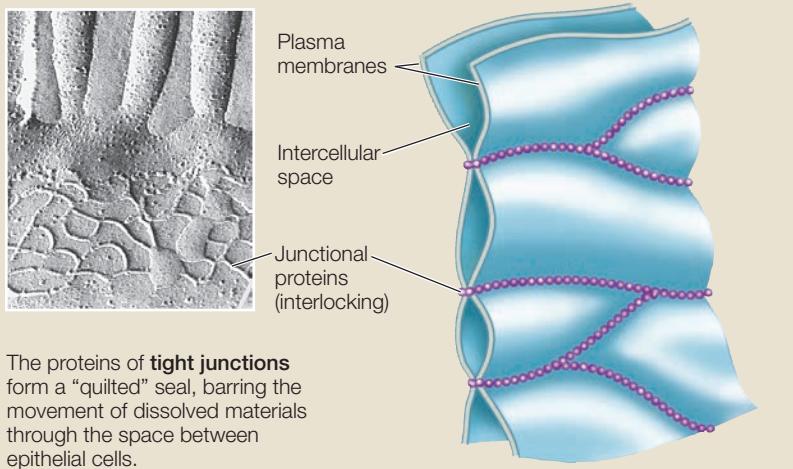
sponge cells). As we describe in Section 3.2, a protein not only has a specific shape, but also has specific chemical groups exposed on its surface where they can interact with other substances, including other proteins. Both of these features allow binding to other specific molecules. The cells of the disaggregated sponge in Figure 6.6 find one another again through the recognition of exposed chemical groups on their membrane glycoproteins. Adhesion proteins are not restricted to animal cells. In most plant cells, the plasma membrane is covered with a thick cell wall, but this structure also has adhesion proteins that allow cells to bind to one another.

In most cases, the binding of cells in a tissue is **homotypic**; that is, the same molecule sticks out of both cells, and the exposed surfaces bind to each other. But **heterotypic** binding (of cells with different proteins) can also occur. In this case, different chemical groups on different surface molecules have an affinity for one another. For example, when the mammalian sperm meets the egg, different proteins on the two types of cells have complementary binding surfaces. Similarly, some algae form male and female reproductive cells (analogous to sperm and eggs) that have flagella to propel them toward each other. Male and female cells can recognize each other by heterotypic proteins on their flagella.

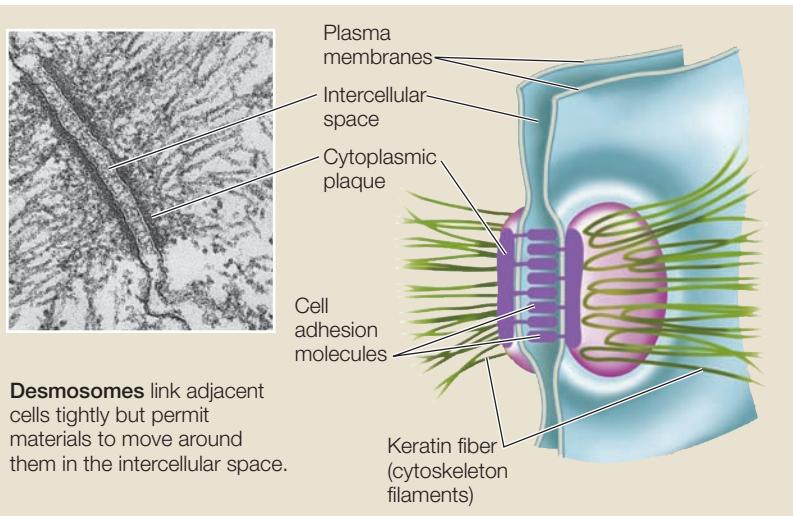
Three types of cell junctions connect adjacent cells

In a complex multicellular organism, cell recognition proteins allow specific types of cells to bind to one another. Often, after

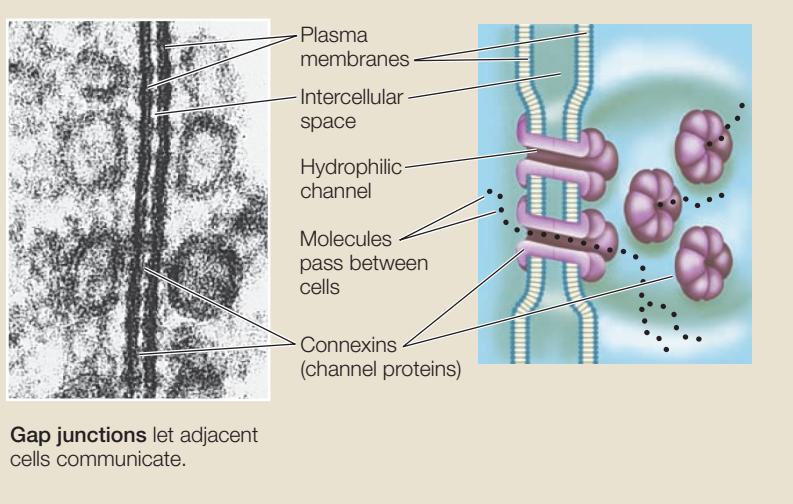
(A)



(B)



(C)

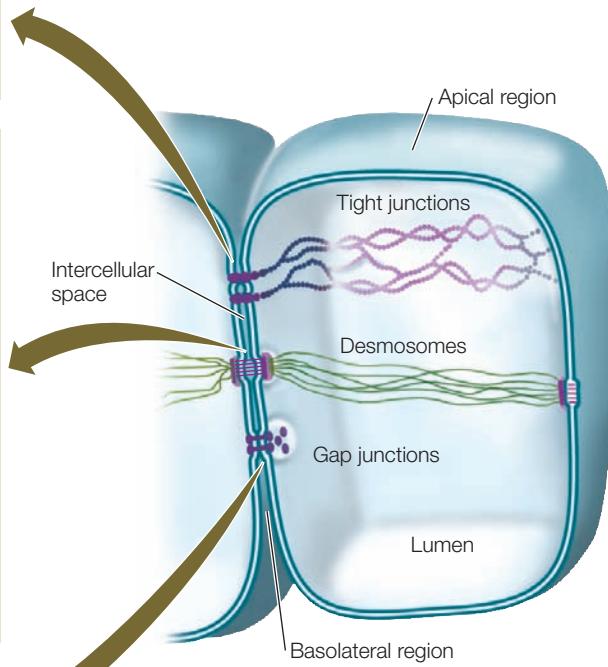


the initial binding, both cells contribute material to form additional membrane structures that connect them to one another. These specialized structures, called **cell junctions**, are most evident in electron micrographs of *epithelial* tissues, which are layers of cells that line body cavities or cover body surfaces. These surfaces often receive stresses, or must retain contents under

6.7 Junctions Link Animal Cells Together Tight junctions (A) and desmosomes (B) are abundant in epithelial tissues. Gap junctions (C) are also found in some muscle and nerve tissues, in which rapid communication between cells is important. Although all three junction types are shown in the cell at the right, all three are not necessarily seen at the same time in actual cells.

yourBioPortal.com

GO TO Web Activity 6.2 • Animal Cell Junctions



pressure, or both, so it is particularly important that their cells adhere tightly. We will examine three types of cell junctions that enable animal cells to seal intercellular spaces, reinforce attachments to one another, and communicate with each other. Tight junctions, desmosomes, and gap junctions, respectively, perform these three functions.

TIGHT JUNCTIONS SEAL TISSUES **Tight junctions** are specialized structures that link adjacent epithelial cells, and they result from the mutual binding of specific proteins in the plasma membranes of the cells. These proteins are arrayed in bands so that they form a series of joints encircling each cell (**Figure 6.7A**). Tight junctions are found in the lining of lumens (cavities) in organs such as the stomach and intestine. They have two major functions:

- They prevent substances from moving from the lumen through the spaces between cells. For example, the presence of tight junctions means that substances must pass through, rather than between, the epithelial cells that form

the lining of the digestive tract. In another example, the cells lining the bladder have tight junctions so urine cannot leak out into the body cavity. Thus, tight junctions help to establish cellular control over what enters and leaves the body.

- They define specific functional regions of membranes by restricting the migration of membrane proteins and phospholipids from one region of the cell to another. Thus the membrane proteins and phospholipids in the apical (“tip”) region of an intestinal epithelial cell (facing the lumen) are different from those in the basolateral (*basal*, “bottom”; *lateral*, “side”) regions of the cell (facing the body cavity or blood capillary outside the lumen).

By forcing materials to enter certain cells, and by allowing different areas of the same cell to have different membrane proteins with different functions, tight junctions in the digestive tract help ensure the directional movement of materials into the body.

DESMOSOMES HOLD CELLS TOGETHER Desmosomes connect adjacent plasma membranes. Desmosomes hold neighboring cells firmly together, acting like spot welds or rivets (Figure 6.7B). Each desmosome has a dense structure called a plaque on the cytoplasmic side of the plasma membrane. To this plaque are attached special cell adhesion molecules that stretch from the plaque through the plasma membrane of one cell, across the intercellular space, and through the plasma membrane of the adjacent cell, where they bind to the plaque proteins in that adjacent cell.

The plaque is also attached to fibers in the cytoplasm. These fibers, which are intermediate filaments of the cytoskeleton (see Figure 5.18), are made of a protein called keratin. They stretch from one cytoplasmic plaque across the cell to another plaque on the other side of the cell. Anchored thus on both sides of the cell, these extremely strong fibers provide great mechanical stability to epithelial tissues. This stability is needed for these tissues, which often receive rough wear while protecting the integrity of the organism’s body surface, or the surface of an organ.

GAP JUNCTIONS ARE A MEANS OF COMMUNICATION Whereas tight junctions and desmosomes have mechanical roles, gap junctions facilitate communication between cells. Each gap junction is made up of specialized channel proteins, called *connexins*, which interact to form a structure (called a *connexon*) that spans the plasma membranes of adjacent cells and the intercellular space between them (Figure 6.7C). Water, dissolved small molecules, and ions can pass from cell to cell through these junctions. This allows groups of cells to coordinate their activities. In Chapter 7 we discuss cell communication and signaling, and in that chapter we describe in more detail the roles of gap junctions and plasmodesmata, which perform a similar role in plants.

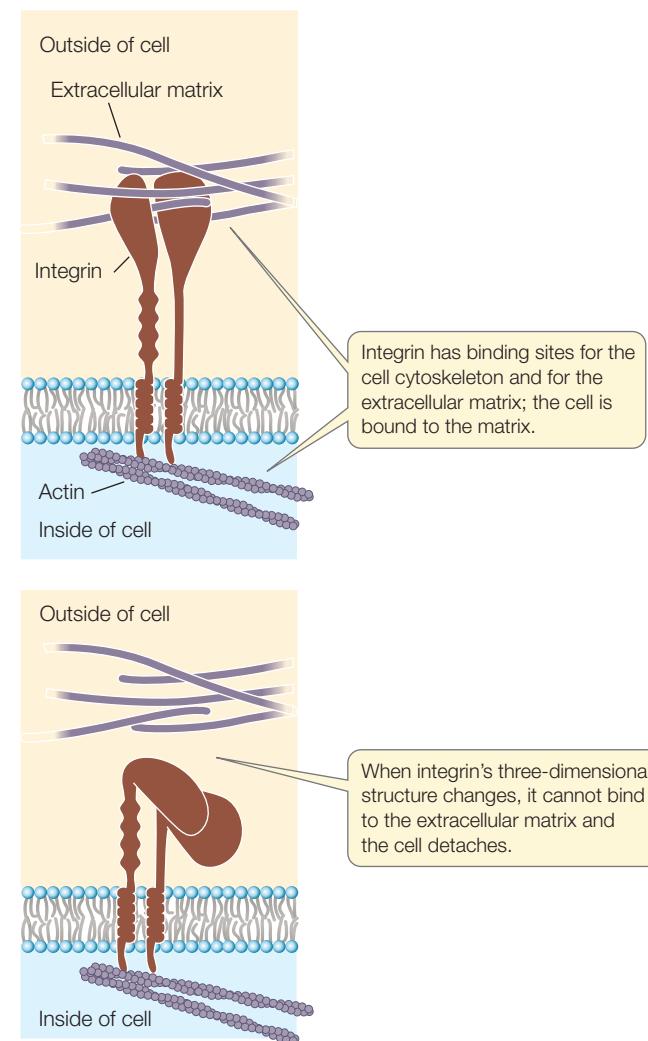
Cell membranes adhere to the extracellular matrix

In Section 5.4 we describe the extracellular matrix of animal cells, which is composed of collagen protein arranged in fibers

in a gelatinous matrix of proteoglycans. The attachment of a cell to the extracellular matrix is important in maintaining the integrity of a tissue. In addition, some cells can detach from their neighbors, move, and attach to other cells; this is often mediated by interactions with the extracellular matrix.

A transmembrane protein called **integrin** often mediates the attachment of epithelial cells to the extracellular matrix (Figure 6.8). More than 24 different integrins have been described in human cells. All of them bind to a protein in the extracellular matrix on the outside of the cell, and to actin filaments, which are part of the cytoskeleton, on the inside of the cell. So, in addition to adhesion, integrin has a role in maintaining cell structure via its interaction with the cytoskeleton.

The binding of integrin to the extracellular matrix is noncovalent and reversible. When a cell moves its location within a tissue or organism, the first step is detachment of the cell’s integrin from the matrix. The integrin protein changes its three-dimensional structure and no longer maintains its link to the matrix. These events are important for cell movement within the developing embryo, and for the spread of cancer cells.



6.8 Integrins Mediate the Attachment of Animal Cells to the Extracellular Matrix

6.2 RECAP

In multicellular organisms, cells arrange themselves in groups by two processes: cell recognition and cell adhesion. Both processes are mediated by integral proteins in the plasma membrane. Cell membrane proteins also interact with the extracellular matrix.

- Describe the difference between cell recognition and cell adhesion. **See p. 111**
- How do the three types of cell junctions regulate the passage of materials between cells and through the intercellular space? **See pp. 111–113 and Figure 6.7**

We have just examined how the plasma membrane structure accommodates the binding and maintenance of cell adhesion. We turn now to another major function of membranes: regulating the substances that enter or leave a cell or organelle.

6.3 What Are the Passive Processes of Membrane Transport?

As you have already learned, biological membranes have many functions, and control of the cell's internal composition is one of the most significant. Biological membranes allow some substances, but not others, to pass through them. This characteristic of membranes is called **selective permeability**. Selective permeability allows the membrane to determine what substances enter or leave a cell or organelle.

There are two fundamentally different processes by which substances cross biological membranes:

- The processes of **passive transport** do not require any input of outside energy to drive them (no metabolic energy).
- The processes of **active transport** require the input of chemical energy from an outside source (metabolic energy).

This section focuses on the passive processes by which substances cross membranes. The energy for the passive transport of a substance is found in the difference between its concentration on one side of the membrane and its concentration on the other. Passive transport processes include two types of diffusion: simple diffusion through the phospholipid bilayer, and facilitated diffusion through *channel proteins* or by means of *carrier proteins*.

Diffusion is the process of random movement toward a state of equilibrium

Nothing in this world is ever absolutely at rest. Everything is in motion, although the motions may be very small. An important consequence of all this random vibration, rotation and translocation (moving from one location to another) of molecules is that all the components of a solution tend eventually to become evenly distributed. For example, if a drop of ink is allowed to

fall into a container of water, the pigment molecules of the ink are initially very concentrated. Without human intervention, such as stirring, the pigment molecules move about at random, spreading slowly through the water until eventually the concentration of pigment—and thus the intensity of color—is exactly the same in every drop of liquid in the container.

A solution in which the solute particles are uniformly distributed is said to be at *equilibrium* because there will be no future net change in their concentration. Equilibrium does not mean that the particles have stopped moving; it just means that they are moving in such a way that their overall distribution does not change.

Diffusion is the process of random movement toward a state of equilibrium. Although the motion of each individual particle is absolutely random, the net movement of particles is directional until equilibrium is reached. Diffusion is thus a net movement from regions of greater concentration to regions of lesser concentration (**Figure 6.9**).

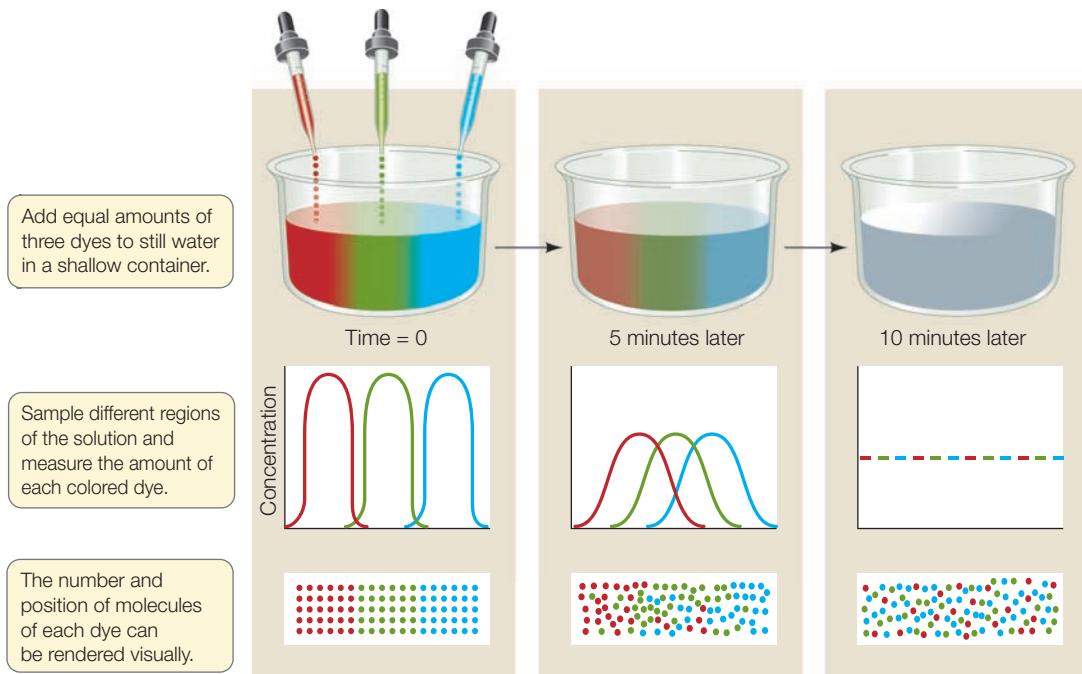
In a complex solution (one with many different solutes), the diffusion of each solute is independent of those of the others. How fast a substance diffuses depends on three factors:

- The *diameter* of the molecules or ions: smaller molecules diffuse faster.
- The *temperature* of the solution: higher temperatures lead to faster diffusion because ions or molecules have more energy, and thus move more rapidly, at higher temperatures.
- The *concentration gradient* in the system—that is, the change in solute concentration with distance in a given direction: the greater the concentration gradient, the more rapidly a substance diffuses.

We'll see how these factors influence membrane transport in the detailed discussions that follow.

DIFFUSION WITHIN CELLS AND TISSUES Within cells, or wherever distances are very short, solutes distribute themselves rapidly by diffusion. Small molecules and ions may move from one end of an organelle to another in a millisecond (10^{-3} s, or one-thousandth of a second). However, the usefulness of diffusion as a transport mechanism declines drastically as distances become greater. In the absence of mechanical stirring, diffusion across more than a centimeter may take an hour or more, and diffusion across meters may take years! Diffusion would not be adequate to distribute materials over the length of a human body, much less that of a larger organism. But within our cells or across layers of one or two cells, diffusion is rapid enough to distribute small molecules and ions almost instantaneously.

DIFFUSION ACROSS MEMBRANES In a solution without barriers, all the solutes diffuse at rates determined by temperature, their physical properties, and their concentration gradients. If a biological membrane divides the solution into separate compartments, then the movement of the different solutes can be affected by the properties of the membrane. The membrane is said to be *permeable* to solutes that can cross it more or less easily, but *impermeable* to substances that cannot move across it.



6.9 Diffusion Leads to Uniform Distribution of Solutes

A simple experiment demonstrates that solutes move from regions of greater concentration to regions of lesser concentration until equilibrium is reached.

Molecules to which the membrane is impermeable remain in separate compartments, and their concentrations may be different on the two sides of the membrane. Molecules to which the membrane is permeable diffuse from one compartment to the other until their concentrations are equal on both sides of the membrane. When the concentrations of a diffusing substance on the two sides of the permeable membrane are identical, equilibrium is reached. Individual molecules continue to pass through the membrane after equilibrium is established, but equal numbers of molecules move in each direction, so *at equilibrium there is no net change in concentration*.

Simple diffusion takes place through the phospholipid bilayer

In **simple diffusion**, small molecules pass through the phospholipid bilayer of the membrane. A molecule that is itself hydrophobic, and is therefore soluble in lipids, enters the membrane readily and is able to pass through it. The more lipid-soluble the molecule is, the more rapidly it diffuses through the membrane bilayer. This statement holds true over a wide range of molecular weights.

On the other hand, electrically charged or polar molecules, such as amino acids, sugars, and ions, do not pass readily through a membrane for two reasons. First, such charged or polar molecules are not very soluble in the hydrophobic interior of the bilayer. Second, such charged and polar substances form many hydrogen bonds with water and ions in the aqueous environment, be it the cytoplasm or the cell exterior. The multiplicity of these hydrogen bonds prevent the substances from moving into the hydrophobic interior of the membrane.

Consider two molecules: a small protein made up of a few polar amino acids, and a cholesterol-based steroid of equivalent size. If a membrane separates high and low concentrations of these substances, the protein, being polar, will diffuse only very

slowly through the membrane, while the nonpolar steroid will diffuse through it readily.

Osmosis is the diffusion of water across membranes

Water molecules pass through specialized channels in membranes (see below) by a diffusion process called **osmosis**. This completely passive process uses no metabolic energy and can be understood in terms of solute concentrations. Recall that a solute dissolves in a solvent and the solute's constituents are dispersed throughout the solution. Osmosis depends on the *number* of solute particles present, not on the *kinds* of particles. We will describe osmosis using red blood cells and plant cells as examples. In these examples, the plasma membranes are considered to be permeable to water and impermeable to most solutes.

Red blood cells are normally suspended in a fluid called plasma, which contains salts, proteins, and other solutes. Examining a drop of blood under the light microscope reveals that these red cells have a characteristic flattened disk shape with a depressed center, sometimes called "biconcave." If pure water is added to the drop of blood, drastically reducing the solute concentration of the plasma, the red cells quickly swell and burst. Similarly, if slightly wilted lettuce is placed in pure water, it soon becomes crisp; by weighing it before and after, we can show that it has taken up water. If, on the other hand, red blood cells or crisp lettuce leaves are placed in a relatively concentrated solution of salt or sugar, the leaves become limp (they wilt), and the red blood cells pucker and shrink.

From such observations we know that the difference in solute concentration between a cell and its surrounding environment determines whether water will move from the environment into the cell or out of the cell into the environment. Other things being equal, if two different solutions are separated by a membrane that allows water, *but not solutes*, to pass through, water molecules will move across the membrane toward the solu-

tion with a higher solute concentration. In other words, water will diffuse from a region of its higher concentration (with a lower concentration of solutes) to a region of its lower concentration (with a higher concentration of solutes).

Three terms are used to compare the solute concentrations of two solutions separated by a membrane:

- A **hypertonic** solution has a higher solute concentration than the other solution with which it is being compared (**Figure 6.10A**).
- **Isotonic** solutions have equal solute concentrations (**Figure 6.10B**).
- A **hypotonic** solution has a lower solute concentration than the other solution with which it is being compared (**Figure 6.10C**).

Water moves from a hypotonic solution across a membrane to a hypertonic solution.

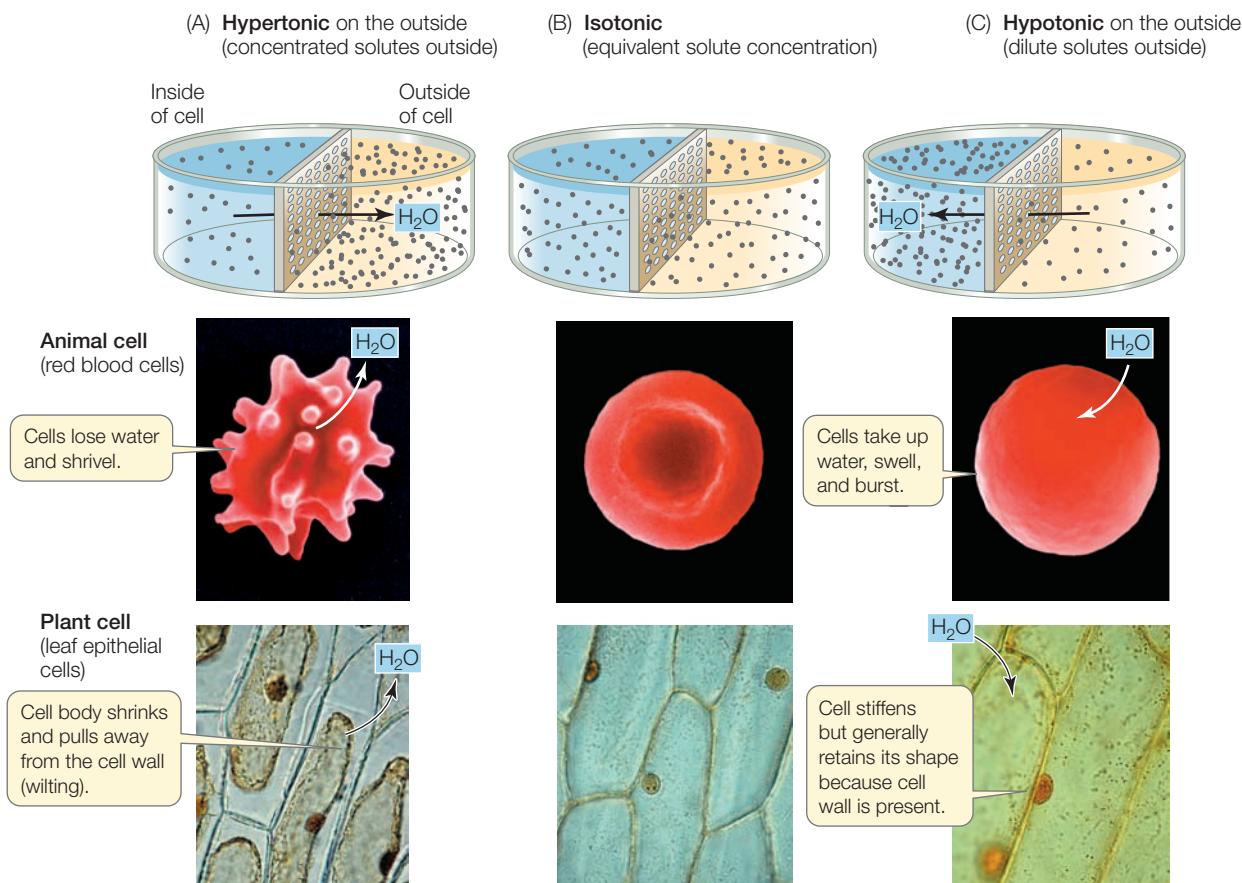
When we say that “water moves,” bear in mind that we are referring to the net movement of water. Since it is so abundant, water is constantly moving through protein channels across the plasma membrane into and out of cells. What concerns us here is whether the overall movement is greater in one direction or the other.

The concentration of solutes in the environment determines the direction of osmosis in all animal cells. A red blood cell takes up water from a solution that is hypotonic to the cell’s contents.

The cell bursts because its plasma membrane cannot withstand the pressure created by the water entry and the resultant swelling. The integrity of red blood cells (and other blood cells) is absolutely dependent on the maintenance of a constant solute concentration in the blood plasma: the plasma must be isotonic to the blood cells if the cells are not to burst or shrink. Regulation of the solute concentration of body fluids is thus an important process for organisms without cell walls.

In contrast to animal cells, the cells of plants, archaea, bacteria, fungi, and some protists have cell walls that limit their volumes and keep them from bursting. Cells with sturdy walls take up a limited amount of water, and in so doing they build up internal pressure against the cell wall, which prevents further water from entering. This pressure within the cell is called **turgor pressure**. Turgor pressure keeps plants upright (and lettuce crisp) and is the driving force for the enlargement of plant cells. It is a normal and essential component of plant growth. If enough water leaves the cells, turgor pressure drops and the plant wilts. Turgor pressure reaches about 100 pounds per square inch (0.7 kg/cm^2)—several times greater than the pres-

6.10 Osmosis Can Modify the Shapes of Cells In a solution that is isotonic with the cytoplasm (center column), a plant or animal cell maintains a consistent, characteristic shape because there is no net movement of water into or out of the cell. In a solution that is hypotonic to the cytoplasm (right), water enters the cell. An environment that is hypertonic to the cytoplasm (left) draws water out of the cell.



sure in automobile tires. This pressure is so great that the cells would change shape and detach from one another, were it not for adhesive molecules in the plant cell wall.

Diffusion may be aided by channel proteins

As we saw earlier, polar or charged substances such as water, amino acids, sugars and ions do not readily diffuse across membranes. But they can cross the hydrophobic phospholipid bilayer passively (that is, without the input of energy) in one of two ways, depending on the substance:

- **Channel proteins** are integral membrane proteins that form channels across the membrane through which certain substances can pass.
- Some substances can bind to membrane proteins called **carrier proteins** that speed up their diffusion through the phospholipid bilayer.

Both of these processes are forms of **facilitated diffusion**. That is, the substances diffuse according to their concentration gradients, but their diffusion is facilitated by protein channels or carriers.

ION CHANNELS The best-studied channel proteins are the **ion channels**. As you will see in later chapters, the movement of ions across membranes is important in many biological processes, ranging from respiration within the mitochondria, to the electrical activity of the nervous system and the opening of the pores in leaves that allow gas exchange with the environment. Several types of ion channels have been identified, each of them specific for a particular ion. All of them show the same basic structure of a hydrophilic pore that allows a particular ion to move through it (**Figure 6.11**).

Just as a fence may have a gate that can be opened or closed, most ion channels are gated: they can be opened or closed to ion passage. A **gated channel** opens when a stimulus causes a change in the three-dimensional shape of the channel. In some cases, this stimulus is the binding of a chemical signal, or **ligand** (see Figure 6.11). Channels controlled in this way are called *ligand-gated channels*. In contrast, a *voltage-gated channel* is stimulated to open or close by a change in the voltage (electrical charge difference) across the membrane.

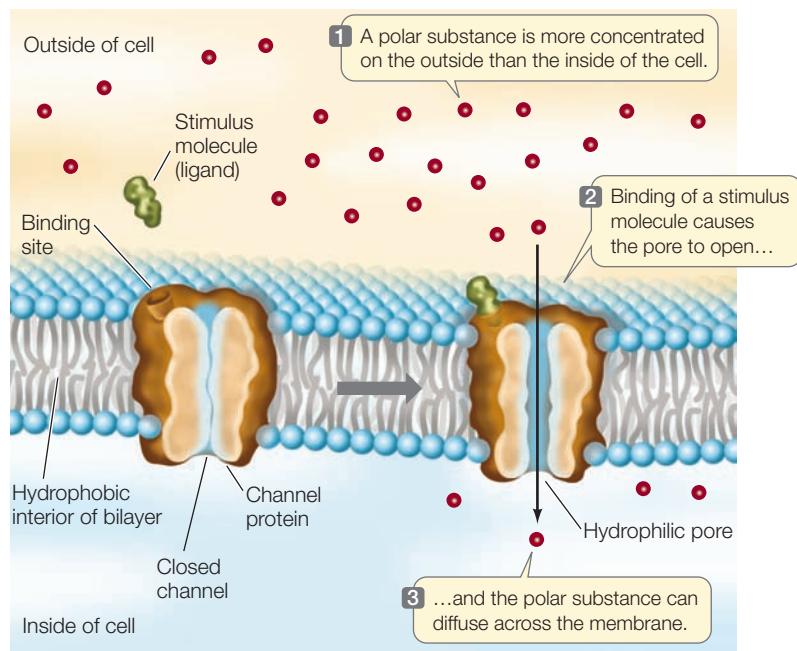
THE MEMBRANE POTENTIAL All living cells maintain an imbalance of ion concentrations across the plasma membrane, and consequently a small voltage or **membrane potential** exists across that membrane. When a gated ion

6.11 A Gated Channel Protein Opens in Response to a Stimulus The channel protein has a pore of polar amino acids and water. It is anchored in the hydrophobic bilayer interior by its outer coating of nonpolar R groups of its amino acids. The protein changes its three-dimensional shape when a stimulus molecule (ligand) binds to it, opening the pore so that hydrophilic polar substances can pass through. Other gated channels open in response to an electrical potential (voltage).

channel opens, millions of ions can rush through it per second. How fast the ions move, and in which direction (into or out of the cell), depends on two factors, the concentration gradient and the magnitude of the voltage. Let's consider how these factors affect the concentration of potassium ions (K^+) inside an animal cell:

- *The concentration gradient:* Because of active transport (discussed below), the concentration of K^+ is usually much higher inside the cell than outside, so K^+ will tend to diffuse out of the cell through an open potassium channel.
- *The distribution of electrical charge:* As K^+ diffuses out of the cell it leaves behind an excess of chloride (Cl^-) and other negatively charged ions. These negatively charged substances cannot readily diffuse through the plasma membrane to follow K^+ out of the cell, and this results in a charge difference (negative inside) across the membrane. K^+ is attracted to the negative charge inside the cell, creating a tendency for K^+ to stay inside the cell, even though it is more concentrated there than outside.

Now, consider what happens when the K^+ channel is opened. Two forces are at work: diffusion draws K^+ out of the cell through the channel, and electrical attraction keeps K^+ inside the cell. The system exists in a state of equilibrium, in which the ion's rate of diffusion out through the channel is balanced by the rate of movement in through the channel due to electrical attraction. Obviously, the concentrations of K^+ on each side of the membrane will not be equal, as we would expect if diffusion were the only force involved. Instead, the attraction of electrical charges keeps some extra K^+ inside the cell. This imbalance in K^+ is a major factor in generating a voltage across the plasma membrane called the *membrane potential*.



The membrane potential is related to the concentration imbalance of K⁺ by the Nernst equation:

$$E_K = 2.3 \frac{RT}{zF} \log \left[\frac{[K]_o}{[K]_i} \right]$$

where R is the gas constant, F is the Faraday constant (both familiar to chemistry students), T is the temperature, and z is the charge on the ion (+1). Solving for 2.3 RT/zF at 20°C ("room temperature"), the equation becomes much simpler:

$$E_K = 58 \log \left[\frac{[K]_o}{[K]_i} \right]$$

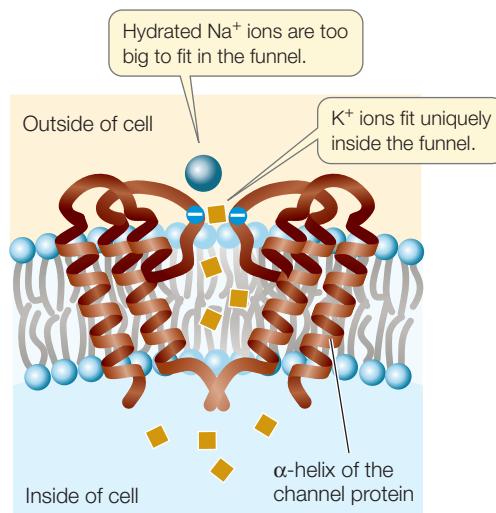
where E_K is the membrane potential (in millivolts, mV) that results from the ratio of K⁺ concentrations outside the cell [K]_o and inside the cell [K]_i.

What does this equation tell us about cells? It shows that a small change in K⁺ concentration, due to the opening of a ligand-gated K⁺ channel, for example, can have a large effect on the electrical potential (E) across the membrane. This change in potential might be enough to cause other proteins in the membrane, such as voltage-gated channels, to change configuration. As we discuss in Chapter 45, this is exactly what happens in the nervous system. Many drugs that act on electrically sensitive tissues work as ligands that open ion channels and thereby affect membrane potential. And as you will see shortly, membrane potential drives secondary active transport.

Actual measurements from animal cells give a total membrane potential between -60 and -70 mV across the membrane, where the inside is negative with respect to the outside (see Figure 45.5). Cells have a tremendous amount of potential energy stored in their membrane potentials. In fact, the brain cells you are using to read this book have more potential energy—about 200,000 volts per centimeter—than the high-voltage electric lines powering your reading light, which carry about 2 volts per centimeter.

THE SPECIFICITY OF ION CHANNELS How does an ion channel allow one ion, but not another, to pass through? It is not simply a matter of charge and size of the ion. For example, a sodium ion (Na⁺), with a radius of 0.095 nanometers, is smaller than K⁺ (0.130 nm), and both carry the same positive charge. Yet the potassium channel lets only K⁺ pass through the membrane, and not the smaller Na⁺. Nobel Laureate Roderick MacKinnon at The Rockefeller University found an elegant explanation for this when he deciphered the structure of a potassium channel from a bacterium (**Figure 6.12**).

Being charged, both Na⁺ and K⁺ are attracted to water molecules. They are surrounded by water "shells" in solution, held by the attraction of their positive charges to the negatively charged oxygen atoms on the water molecules (see Figure 2.10). The potassium channel contains highly polar oxygen atoms at its opening. The gap enclosed by these atoms is exactly the right size so that when a K⁺ ion approaches the opening, it is more strongly attracted to the oxygen atoms there than to those of the



6.12 The Potassium Channel The positively charged potassium ions are attracted by the polar (negatively charged) oxygen atoms in the R groups (side chains) of the channel protein, and the ions funnel through the channel. This channel is a "custom fit" for K⁺; other ions do not pass through.

water molecules in its shell. It sheds its water shell and passes through the channel. The smaller Na⁺ ion, on the other hand, is kept a bit more distant from the oxygen atoms at the opening of the channel because extra water molecules can fit between the ion (with its shell) and the oxygen atoms at the opening. So Na⁺ does not enter the potassium channel. The gate that opens or closes the channel appears to be an interaction between positively charged arginine residues on the protein and negative charges on membrane phospholipids. This is an example of the functional interactions between membrane proteins and lipids.

AQUAPORINS FOR WATER Water crosses membranes at a much faster rate than would be expected for simple diffusion through the hydrophobic phospholipid bilayer. One way that water can do this is by "hitchhiking" with some ions, such as Na⁺, as they pass through ion channels. Up to 12 water molecules may coat an ion as it traverses a channel. But there is an even faster way to get water across membranes. Plant cells and some animal cells, such as red blood cells and kidney cells, have membrane channels called **aquaporins**. These channels function as a cellular plumbing system for moving water. Like the K⁺ channel, the aquaporin channel is highly specific. Water molecules move in single file through the channel, which excludes ions so that the electrical properties of the cell are maintained.

Aquaporins were first identified by Peter Agre at Duke University, who shared the Nobel Prize with Rod MacKinnon (see above). Agre noticed a membrane protein that was present in red blood cells, kidney cells, and plant cells but did not know its function. A colleague suggested that it might be a water channel, because these cell types show rapid diffusion of water across their membranes. Agre inserted the protein into the membrane of an oocyte, which normally does not permit much diffusion of water. He injected the oocyte with mRNA for aquaporin, from which the protein was produced and inserted into

the membrane. Remarkably, the oocyte began swelling immediately after being transferred to a hypotonic solution, indicating rapid diffusion of water into the cell (**Figure 6.13**).

Carrier proteins aid diffusion by binding substances

As we described earlier, another kind of facilitated diffusion involves not just the opening of a channel, but also the actual binding of the transported substance to a membrane protein called a carrier protein. Like channel proteins, carrier proteins allow diffusion both into and out of the cell or organelle. In other words, carrier proteins operate in both directions. Carrier proteins transport polar molecules such as sugars and amino acids.

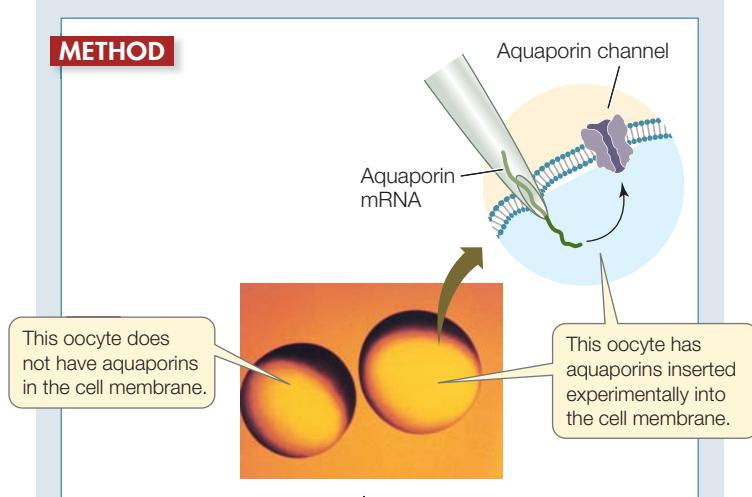
INVESTIGATING LIFE

6.13 Aquaporin Increases Membrane Permeability to Water

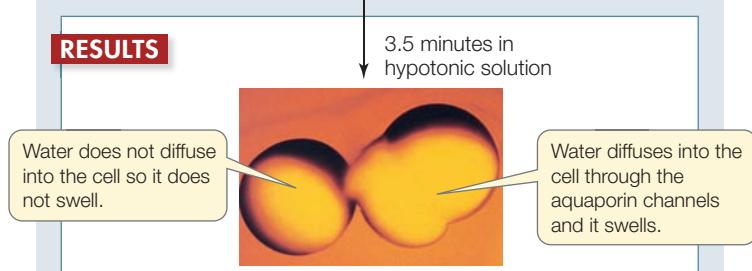
A protein was isolated from the membranes of cells in which water diffuses rapidly across the membranes. When the protein was inserted into oocytes, which do not normally have it, the water permeability of the oocytes was greatly increased.

HYPOTHESIS Aquaporin increases membrane permeability to water.

METHOD



RESULTS



CONCLUSION Aquaporin increases the rate of water diffusion across the cell membrane.

Go to yourBioPortal.com for original citations, discussions, and relevant links for all INVESTIGATING LIFE figures.

Glucose is the major energy source for most mammalian cells, and they require a great deal of it. Their membranes contain a carrier protein—the glucose transporter—that facilitates glucose uptake into the cell. Binding of glucose to a specific three-dimensional site on one side of the transporter protein causes the protein to change its shape and release glucose on the other side of the membrane (**Figure 6.14A**). Since glucose is broken down almost as soon as it enters a cell, there is almost always a strong concentration gradient favoring glucose entry (that is, a higher concentration outside the cell than inside). The transporter allows glucose molecules to cross the membrane and enter the cell much faster than they would by simple diffusion through the bilayer. This rapid entry is necessary to ensure that the cell receives enough glucose for its energy needs.

Transport by carrier proteins is different from simple diffusion. In both processes, the rate of movement depends on the concentration gradient across the membrane. However, in carrier-mediated transport, a point is reached at which increases in the concentration gradient are not accompanied by an increased rate of diffusion. At this point, the facilitated diffusion system is said to be *saturated* (**Figure 6.14B**). Because there are only a limited number of carrier protein molecules per unit of membrane area, the rate of diffusion reaches a maximum when all the carrier molecules are fully loaded with solute molecules. Think of waiting for the elevator on the ground floor of a hotel with 50 other people. They can't all get in the elevator (carrier) at once, so the rate of transport (say 10 people at a time) is saturated.

yourBioPortal.com

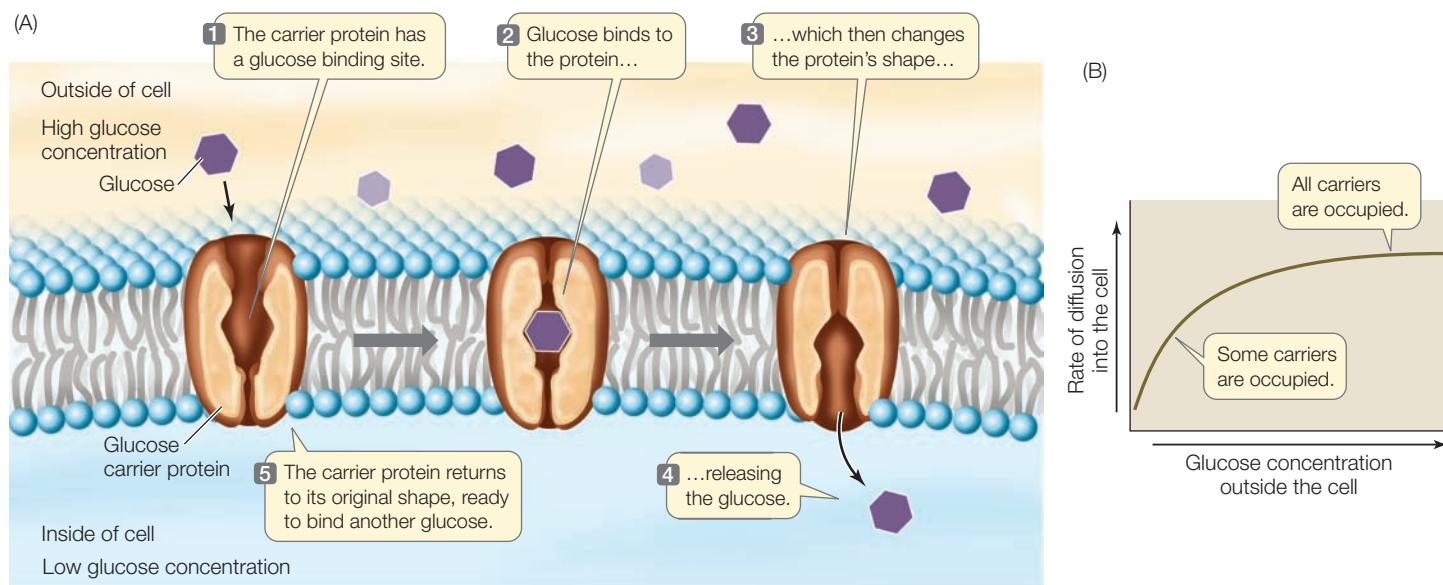
GO TO Animated Tutorial 6.1 • Passive Transport

6.3 RECAP

Diffusion is the movement of ions or molecules from a region of greater concentration to a region of lesser concentration. Water can diffuse through cell membranes by a process called osmosis. Channel proteins, which can be open or closed, and carrier proteins facilitate diffusion of charged and polar substances, including water. The diffusion of ions across cell membranes sets up an electrochemical potential gradient across the membranes.

- What properties of a substance determine whether, and how fast, it will diffuse across a membrane?
See p. 114
- Describe osmosis and explain the terms hypertonic, hypotonic, and isotonic. **See p. 116 and Figure 6.10**
- How does a channel protein facilitate diffusion?
See p. 118 and Figures 6.11 and 6.12

The process of diffusion tends to equalize the concentrations of substances outside and inside cells. However, one hallmark of



6.14 A Carrier Protein Facilitates Diffusion The glucose transporter is a carrier protein that allows glucose to enter the cell at a faster rate than would be possible by simple diffusion. (A) The transporter binds to glucose, brings it into the membrane interior, then changes shape, releasing glucose into the cell cytoplasm. (B) The graph shows the rate of glucose entry via a carrier versus the concentration of glucose outside the cell. As the glucose concentration increases, the rate of diffusion increases until the point at which all the available transporters are being used (the system is saturated).

a living thing is that it can have an internal composition quite different from that of its environment. To achieve this it must sometimes move substances in opposite directions from the ones in which they would naturally tend to diffuse. That is, substances must sometimes be moved against concentration gradients and/or against the cell's membrane potential (electrical gradient). This process requires work—the input of energy—and is known as *active transport*.

6.4 What are the Active Processes of Membrane Transport?

In many biological situations, there is a different concentration of a particular ion or small molecule inside compared with outside a cell. In these cases, the imbalance is maintained by a pro-

tein in the plasma membrane that moves the substance against its concentration and/or electrical gradient. This is called *active transport*, and because it is acting “against the normal flow,” it requires the expenditure of energy. Often the energy source is adenosine triphosphate (ATP). In eukaryotes, ATP is produced in the mitochondria and has chemical energy stored in its terminal phosphate bond. This energy is released when ATP is converted to adenosine diphosphate (ADP) in a hydrolysis reaction that breaks the terminal phosphate bond. This is one source of energy for active transport. (We give the details of how ATP provides energy to cells in Section 8.2.)

The differences between diffusion and active transport are summarized in **Table 6.1**.

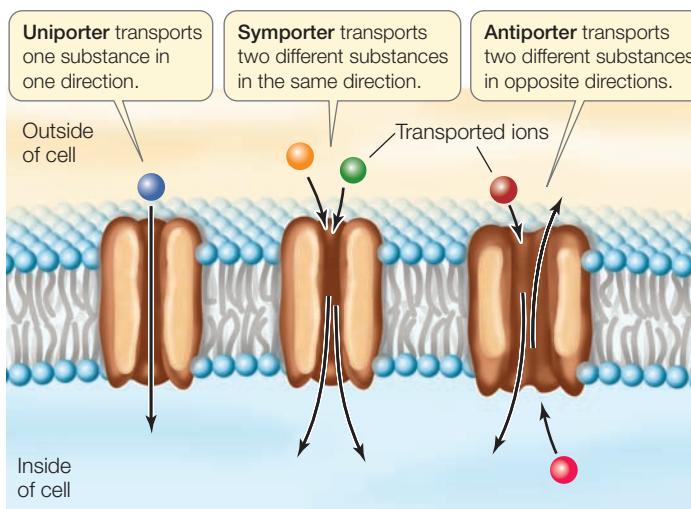
Active transport is directional

Simple and facilitated diffusion follow concentration gradients and can occur in both directions across a membrane. In contrast, active transport is directional, and moves a substance either into or out of the cell or organelle, depending on need. There are three types of active transport, each involving its own type of membrane protein (**Figure 6.15**):

- A **uniporter** moves a single substance in one direction. For example, a calcium-binding protein found in the plasma

TABLE 6.1
Membrane Transport Mechanisms

	SIMPLE DIFFUSION	DIFFUSION THROUGH CHANNEL	FACILITATED DIFFUSION	ACTIVE TRANSPORT
Cellular energy required?	No	No	No	Yes
Driving force	Concentration gradient	Concentration gradient	Concentration gradient	ATP hydrolysis (against concentration gradient)
Membrane protein required?	No	Yes	Yes	Yes
Specificity	No	Yes	Yes	Yes



6.15 Three Types of Proteins for Active Transport Note that in each of the three cases, transport is directional. Symporters and antiporters are examples of coupled transporters. All three types of transporters are coupled to energy sources in order to move substances against their concentration gradients.

membrane and endoplasmic reticulum of many cells actively transports Ca^{2+} to locations where it is more highly concentrated, either outside the cell or inside the ER.

- A **symporter** moves two substances in the same direction. For example, a symporter in the cells that line the intestine must bind Na^+ in addition to an amino acid in order to absorb amino acids from the intestine.

- An **antiporter** moves two substances in opposite directions, one into the cell (or organelle) and the other out of the cell (or organelle). For example, many cells have a sodium-potassium pump that moves Na^+ out of the cell and K^+ into the cell.

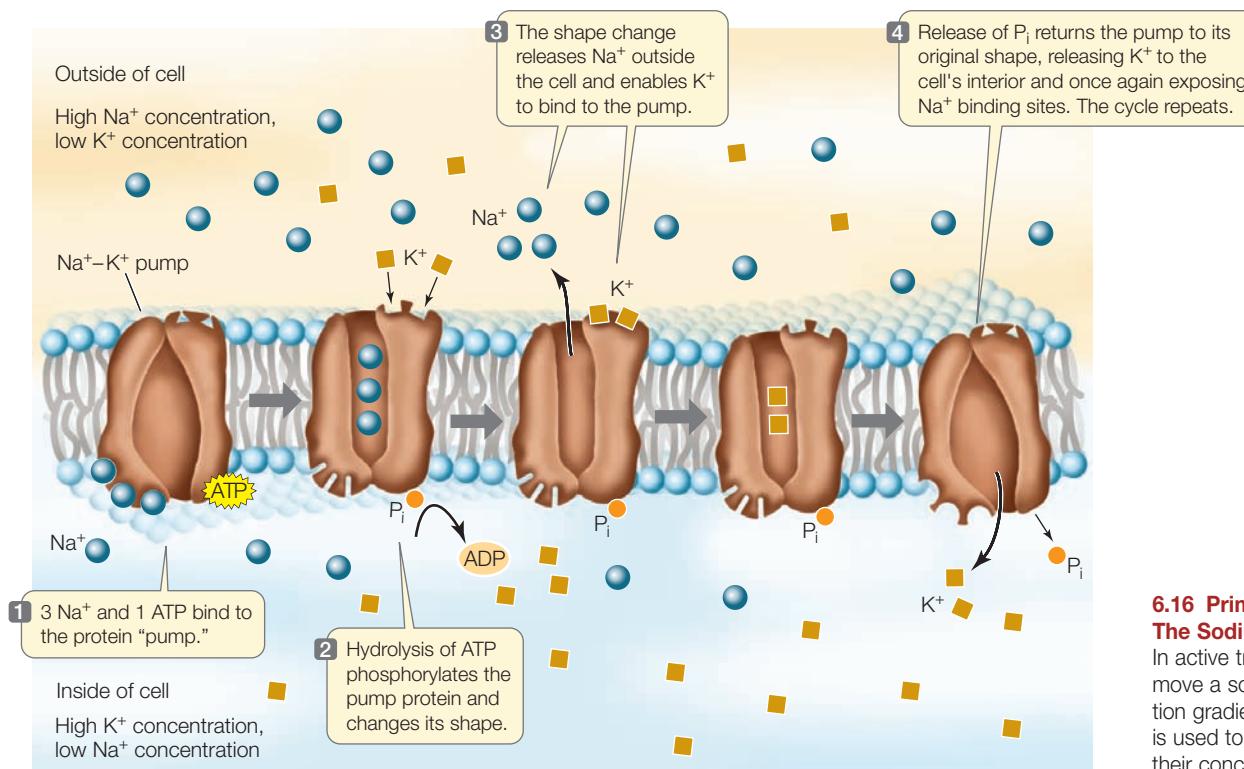
Symporters and antiporters are also known as *coupled transporters* because they move two substances at once.

Different energy sources distinguish different active transport systems

There are two basic types of active transport:

- **Primary active transport** involves the direct hydrolysis of ATP, which provides the energy required for transport.
- **Secondary active transport** does not use ATP directly. Instead, its energy is supplied by an ion concentration and electrical gradient established by primary active transport. This transport system uses the energy of ATP indirectly to set up the gradient.

In primary active transport, energy released by the hydrolysis of ATP drives the movement of specific ions against their concentration gradients. For example, we mentioned earlier that concentrations of potassium ions (K^+) inside a cell are often much higher than in the fluid bathing the cell. On the other hand, the concentration of sodium ions (Na^+) is often much higher outside the cell. A protein in the plasma membrane pumps Na^+ out of the cell and K^+ into the cell against these concentration and electrochemical gradients, ensuring that the gradients are maintained (Figure 6.16). This **sodium-potassium (Na^+-K^+) pump** is



6.16 Primary Active Transport: The Sodium-Potassium Pump

In active transport, energy is used to move a solute against its concentration gradient. Here, energy from ATP is used to move Na^+ and K^+ against their concentration gradients.

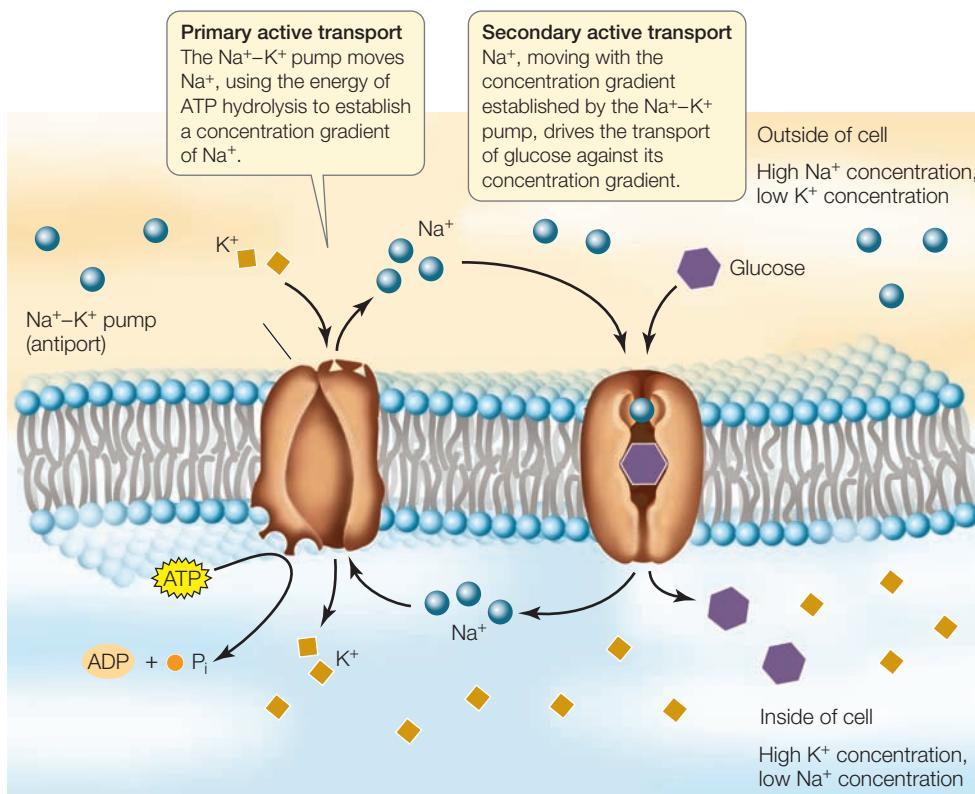
6.17 Secondary Active Transport The Na^+ concentration gradient established by primary active transport (left) powers the secondary active transport of glucose (right). A symporter protein couples the movement of glucose across the membrane against its concentration gradient to the passive movement of Na^+ into the cell.

found in all animal cells. The pump is an integral membrane glycoprotein. It breaks down a molecule of ATP to ADP and a free phosphate ion (P_i) and uses the energy released to bring two K^+ ions into the cell and export three Na^+ ions. The Na^+-K^+ pump is thus an antiporter because it moves two substances in different directions.

In secondary active transport, the movement of a substance against its concentration gradient is accomplished using energy “regained” by letting ions move across the membrane with their electrochemical and concentration gradients. For example, once the sodium–potassium pump establishes a concentration gradient of sodium ions, the passive diffusion of some Na^+ back into the cell can provide energy for the secondary active transport of glucose into the cell (Figure 6.17). This occurs when glucose is absorbed into the bloodstream from the digestive tract. Secondary active transport aids in the uptake of amino acids and sugars, which are essential raw materials for cell maintenance and growth. Both types of coupled transport proteins—symporters and antiporters—are used for secondary active transport.

yourBioPortal.com

GO TO Animated Tutorial 6.2 • Active Transport



We have examined a number of passive and active ways in which ions and small molecules can enter and leave cells. But what about large molecules such as proteins? Many proteins are so large that they diffuse very slowly, and their bulk makes it difficult for them to pass through the phospholipid bilayer. It takes a completely different mechanism to move intact large molecules across membranes.

6.5 How Do Large Molecules Enter and Leave a Cell?

Macromolecules such as proteins, polysaccharides, and nucleic acids are simply too large and too charged or polar to pass through biological membranes. This is actually a fortunate property—think of the consequences if such molecules diffused out of cells. A red blood cell would not retain its hemoglobin! Indeed, as we discuss in Chapter 5, the development of a selectively permeable membrane was essential for the functioning of the first cells when life on Earth began. The interior of a cell can be maintained as a separate compartment with a different composition from that of the exterior environment, which is subject to abrupt changes. On the other hand, cells must sometimes take up or secrete (release to the external environment) intact large molecules. In Section 5.3 we describe phagocytosis, the mechanism by which solid particles can be brought into the cell by means of vesicles that pinch off from the plasma membrane. The general terms for the mechanisms by which substances enter and leave the cell via membrane vesicles are *endocytosis* and *exocytosis*.

6.4 RECAP

Active transport across a membrane is directional and requires an input of energy to move substances against their concentration gradients. Active transport allows a cell to maintain small molecules and ions at concentrations very different from those in the surrounding environment.

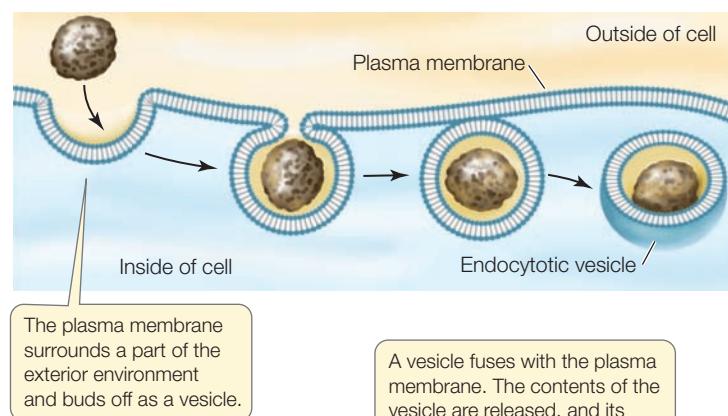
- Why is energy required for active transport? See p. 120
- Explain the difference between primary active transport and secondary active transport. See p. 121
- Why is the sodium–potassium (Na^+-K^+) pump classified as an antiporter? See p. 122 and Figure 6.16

Macromolecules and particles enter the cell by endocytosis

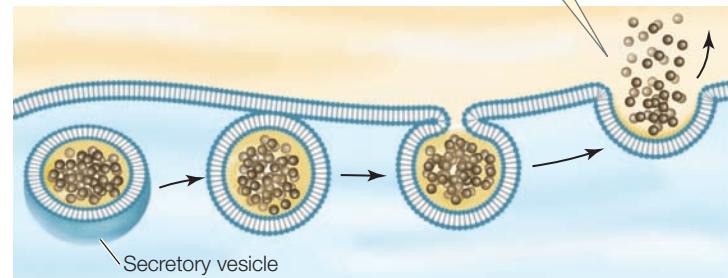
Endocytosis is a general term for a group of processes that bring small molecules, macromolecules, large particles, and even small cells into the eukaryotic cell (**Figure 6.18A**). There are three types of endocytosis: phagocytosis, pinocytosis, and receptor-mediated endocytosis. In all three, the plasma membrane invaginates (folds inward), forming a small pocket around materials from the environment. The pocket deepens, forming a vesicle. This vesicle separates from the plasma membrane and migrates with its contents to the cell's interior.

- In **phagocytosis** ("cellular eating"), part of the plasma membrane engulfs large particles or even entire cells. Unicellular protists use phagocytosis for feeding, and some white blood cells use phagocytosis to defend the body by engulfing foreign cells and substances. The food vacuole or phagosome that forms usually fuses with a lysosome, where its contents are digested (see Figure 5.11).
- In **pinocytosis** ("cellular drinking"), vesicles also form. However, these vesicles are smaller, and the process operates to bring dissolved substances, including proteins or fluids, into the cell. Like phagocytosis, pinocytosis can be relatively nonspecific regarding what it brings into the cell. For example, pinocytosis goes on constantly in the endothelium, the single layer of cells that separates a tiny blood capillary from the surrounding tissue. Pinocytosis allows cells of the endothelium to rapidly acquire fluids and dissolved solutes from the blood.

(A) Endocytosis



(B) Exocytosis



- In **receptor-mediated endocytosis**, molecules at the cell surface recognize and trigger the uptake of specific materials. Let's take a closer look at this last process.

Receptor-mediated endocytosis is highly specific

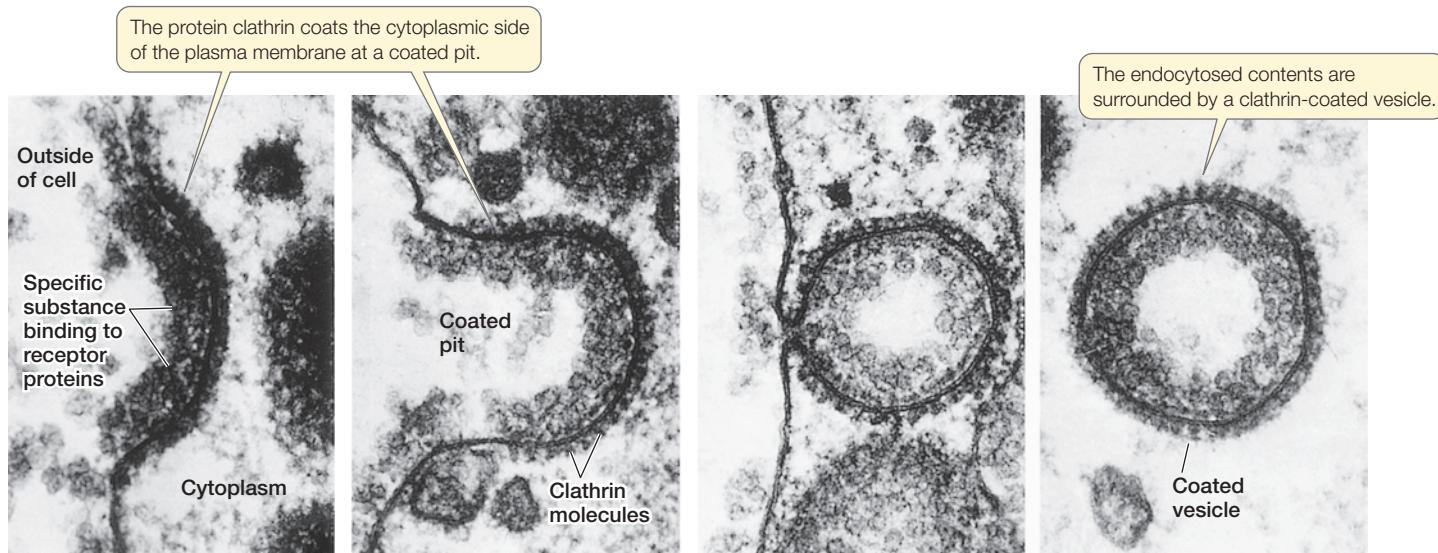
Receptor-mediated endocytosis is used by animal cells to capture specific macromolecules from the cell's environment. This process depends on **receptor proteins**, which are proteins that can bind to specific molecules within the cell or in the cell's external environment. In receptor-mediated endocytosis, the receptors are integral membrane proteins located at particular regions on the extracellular surface of the plasma membrane. These membrane regions are called *coated pits* because they form slight depressions in the plasma membrane and their cytoplasmic surfaces are coated by other proteins, such as clathrin. The uptake process is similar to that in phagocytosis.

When a receptor protein binds to its specific ligand (in this case, the macromolecule to be taken into the cell), its coated pit invaginates and forms a coated vesicle around the bound macromolecule. The clathrin molecules strengthen and stabilize the vesicle, which carries the macromolecule away from the plasma membrane and into the cytoplasm (**Figure 6.19**). Once inside, the vesicle loses its clathrin coat and may fuse with a lysosome, where the engulfed material is digested (by the hydrolysis of polymers to monomers) and the products released into the cytoplasm. Because of its specificity for particular macromolecules, receptor-mediated endocytosis is an efficient method of taking up substances that may exist at low concentrations in the cell's environment.

Receptor-mediated endocytosis is the method by which cholesterol is taken up by most mammalian cells. Water-insoluble cholesterol and triglycerides are packaged by liver cells into lipoprotein particles. Most of the cholesterol is packaged into a type of lipoprotein particle called *low-density lipoprotein*, or LDL, which is circulated via the bloodstream. When a particular cell requires cholesterol, it produces specific LDL receptors, which are inserted into the plasma membrane in clathrin-coated pits. Binding of LDLs to the receptor proteins triggers the uptake of the LDLs via receptor-mediated endocytosis. Within the resulting vesicle, the LDL particles are freed from the receptors. The receptors segregate to a region that buds off and forms a new vesicle, which is recycled to the plasma membrane. The freed LDL particles remain in the original vesicle, which fuses with a lysosome. There, the LDLs are digested and the cholesterol made available for cell use.

In healthy individuals, the liver takes up unused LDLs for recycling. People with the inherited disease *familial hypercholesterolemia* have a deficient LDL receptor in their livers. This prevents receptor-mediated endocytosis of LDLs, resulting in

6.18 Endocytosis and Exocytosis Endocytosis (A) and exocytosis (B) are used by eukaryotic cells to take up and release large molecules and particles, and small cells.



dangerously high levels of cholesterol in the blood. The cholesterol builds up in the arteries that nourish the heart and causes heart attacks. In extreme cases where only the deficient receptor is present, children and teenagers can have severe cardiovascular disease.

Exocytosis moves materials out of the cell

Exocytosis is the process by which materials packaged in vesicles are secreted from a cell when the vesicle membrane fuses with the plasma membrane (see Figure 6.18B). This fusing makes an opening to the outside of the cell. The contents of the vesicle are released into the environment, and the vesicle membrane is smoothly incorporated into the plasma membrane.

In Chapter 5 we encounter exocytosis as the last step in the processing of material engulfed by phagocytosis—the release of undigested materials back to the extracellular environment. Exocytosis is also important in the secretion of many different substances, including digestive enzymes from the pancreas, neurotransmitters from neurons, and materials for the construction of the plant cell wall. You will encounter these processes in later chapters.

yourBioPortal.com

GO TO Animated Tutorial 6.3 • Endocytosis and Exocytosis

6.5 RECAP

Endocytosis and exocytosis are the processes by which large particles and molecules are transported into and out of the cell. Endocytosis may be mediated by a receptor protein in the plasma membrane.

- Explain the difference between phagocytosis and pinocytosis. See p. 123
- Describe an example of receptor-mediated endocytosis. See p. 123 and Figure 6.19

6.19 Receptor-Mediated Endocytosis The receptor proteins in a coated pit bind specific macromolecules, which are then carried into the cell by a coated vesicle.

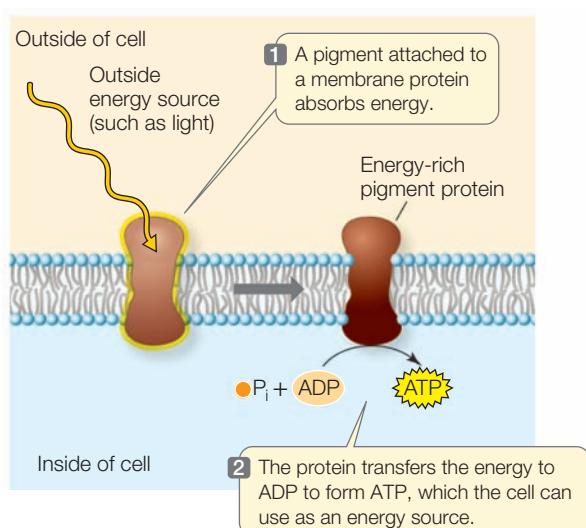
We have now examined the structures and some of the functions of biological membranes. We have seen how macromolecules on the plasma membrane surface allow cells to recognize and adhere to each other, so that tissues and organs can form. We have also seen how membranes selectively regulate the traffic of small and large molecules, and how large particles such as LDLs can be taken up by cells. These are crucial functions, but they are not the only functions of biological membranes.

6.6 What Are Some Other Functions of Membranes?

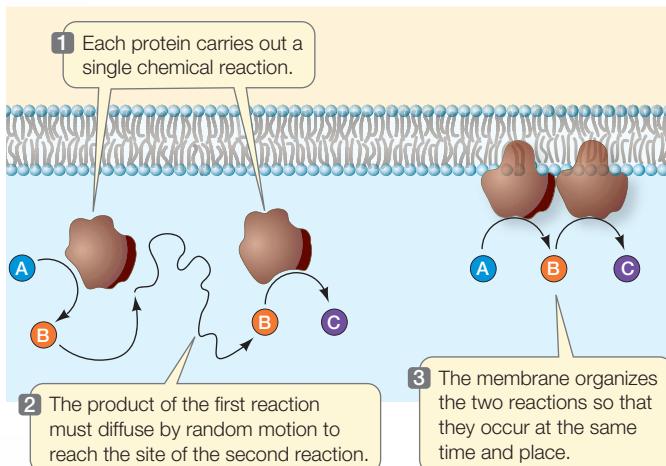
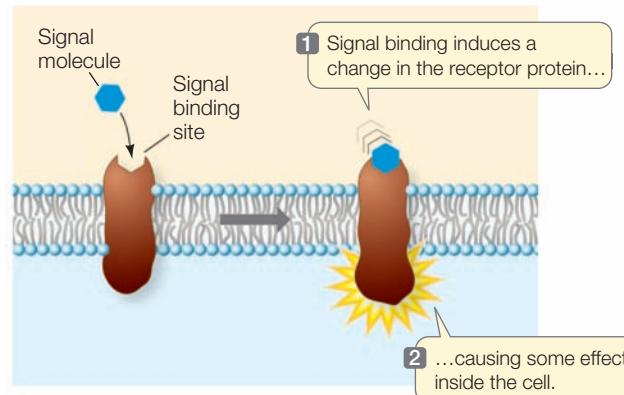
The plasma membranes of certain types of cells, such as neurons and muscle cells, respond to the electric charges carried by ions. These membranes are thus electrically excitable, which gives them important properties. For example, in neurons, the plasma membrane conducts nerve impulses from one end of the cell to the other. In muscle cells, electrical excitation results in muscle contraction.

Other biological activities and properties associated with membranes are discussed in the chapters that follow. Throughout evolution, these activities have been essential for the specialization of cells, tissues, and organisms. Three of these activities are especially important:

- *Some organelle membranes help transform energy (Figure 6.20A).* For example, the inner mitochondrial membrane helps convert the energy of fuel molecules to the energy of phosphate bonds in ATP. The thylakoid membranes of chloroplasts participate in the conversion of light energy to the energy of chemical bonds. These important processes, vital to the life of most eukaryotic organisms, are discussed in detail in Chapters 9 and 10.
- *Some membrane proteins organize chemical reactions.* Often a cellular process depends on a series of enzyme-catalyzed

(A) Energy transformation

6.20 Other Membrane Functions The compartmentation afforded by a lipid bilayer or protein membrane was a key event in the emergence of cells. Functions such as energy transformation (A), organization of chemical reactions (B), and signaling (C) probably evolved later and conferred a selective advantage on cells and organisms that had them.

(B) Organizing chemical reactions**(C) Information processing**

reactions, in which the products of one reaction serve as reactants in the next. For such a series of reactions to occur, all the necessary molecules must come together. In a solution, reactant and enzyme molecules are randomly distributed and collisions among them are random. Because these collisions are necessary for chemical reactions to occur, a complete series of chemical reactions may occur only very slowly in a solution. However, if the different enzymes are

bound to a membrane in sequential order, the product of one reaction can be released close to the enzyme for the next reaction. Such an “assembly line” allows reactions to proceed rapidly and efficiently (Figure 6.20B).

- Some membrane proteins process information. As we have seen, biological membranes may have integral membrane proteins or attached carbohydrates that can bind to specific substances in the environment. Without entering a cell, a specific ligand can bind to a receptor and serve as a signal to initiate, modify, or turn off a cell function (Figure 6.20C). In this type of information processing, specificity in binding is essential.

We have seen the informational role of the LDL receptor protein in the recognition and endocytosis of LDL, with its cargo of cholesterol. Another example is the binding of a hormone such as insulin to specific receptors on a target cell. When insulin binds to receptors on a liver cell, it elicits the uptake of glucose. In Chapter 7 there are many other examples of the role of membrane proteins in information processing.

CHAPTER SUMMARY

6.1 What Is the Structure of a Biological Membrane?

- Biological membranes consist of lipids, proteins, and carbohydrates. The **fluid mosaic model** of membrane structure describes a phospholipid bilayer in which proteins can move about within the plane of the membrane. SEE WEB ACTIVITY 6.1
- The two leaflets of a membrane may have different properties because of their different phospholipid compositions, exposed

domains of **integral membrane proteins**, and **peripheral membrane proteins**. Some proteins, called **transmembrane proteins**, span the membrane. Review Figure 6.1

- Carbohydrates, attached to proteins in **glycoproteins** or to phospholipids in **glycolipids**, project from the external surface of the plasma membrane and function as recognition signals.
- Membranes are not static structures, but are constantly forming, exchanging, and breaking down.

6.2 How Is the Plasma Membrane Involved in Cell Adhesion and Recognition?

- In order for cells to assemble into tissues they must recognize and adhere to one another. **Cell recognition** and **cell adhesion** depend on integral membrane proteins that protrude from the cell surface. Binding can be between the same proteins from two cells (**homotypic**) or different proteins (**heterotypic**).

Review Figure 6.6

- Cell junctions connect adjacent cells. **Tight junctions** prevent the passage of molecules through the intercellular spaces between cells, and they restrict the migration of membrane proteins over the cell surface. **Desmosomes** cause cells to adhere firmly to one another. **Gap junctions** provide channels for communication between adjacent cells. **Review Figure 6.7.**

WEB ACTIVITY 6.2

- Integrins** mediate the attachment of animal cells to the extracellular matrix. **Review Figure 6.8**

6.3 What Are the Passive Processes of Membrane Transport?

SEE ANIMATED TUTORIAL 6.1

- Membranes exhibit **selective permeability**, regulating which substances pass through them.
- A substance can diffuse passively across a membrane by one of two processes: **simple diffusion** through the phospholipid bilayer or **facilitated diffusion** either through a **channel** or by means of a **carrier protein**.
- A solute diffuses across a membrane from a region with a greater concentration of that solute to a region with a lesser concentration of that solute. Equilibrium is reached when the solute concentrations on both sides of the membrane show no net change over time. **Review Figure 6.9**
- In **osmosis**, water diffuses from a region of higher water concentration to a region of lower water concentration.
- Most cells are in an **isotonic** environment, where total solute concentrations on both sides of the plasma membrane are equal. If the solution surrounding a cell is **hypotonic** to the cell interior, more water enters the cell than leaves it. In plant cells, this leads to **turgor pressure**. In a **hypertonic** solution, more water leaves the cell than enters it. **Review Figure 6.10**
- Ion channels** are membrane proteins that allow the rapid facilitated diffusion of ions through membranes. **Gated channels** can be opened or closed by certain conditions or chemicals. The

opening or closing of channels, as well as an asymmetric distribution of charged molecules, sets up an **electrochemical gradient** on different sides of a membrane. **Review Figure 6.11**

- Aquaporins** are water channels. **Review Figure 6.13**
- Carrier proteins** bind to polar molecules such as sugars and amino acids and transport them across the membrane. The maximum rate of this type of facilitated diffusion is limited by the number of carrier (transporter) proteins in the membrane. **Review Figure 6.14**

6.4 What Are the Active Processes of Membrane Transport?

SEE ANIMATED TUTORIAL 6.2

- Active transport** requires the use of chemical energy to move substances across membranes against their concentration gradients. Active transport proteins may be **uniporters**, **symporters**, or **antiporters**. **Review Figure 6.15**
- In **primary active transport**, energy from the hydrolysis of ATP is used to move ions into or out of cells. The **sodium-potassium pump** is an important example. **Review Figure 6.16**
- Secondary active transport** couples the passive movement of one substance down its concentration gradient to the movement of another substance against its concentration gradient. Energy from ATP is used indirectly to establish the concentration gradient that results in the movement of the first substance. **Review Figure 6.17**

6.5 How Do Large Molecules Enter and Leave a Cell?

SEE ANIMATED TUTORIAL 6.3

- Endocytosis** is the transport of macromolecules, large particles, and small cells into eukaryotic cells via the invagination of the plasma membrane and the formation of vesicles. **Phagocytosis** and **pinocytosis** are types of endocytosis. **Review Figure 6.18A**
- In **receptor-mediated endocytosis**, a specific **receptor protein** on the plasma membrane binds to a particular macromolecule.
- In **exocytosis**, materials in vesicles are secreted from the cell when the vesicles fuse with the plasma membrane. **Review Figure 6.18B**

6.6 What Are Some Other Functions of Membranes?

- Membranes function as sites for energy transformations, for organizing chemical reactions, and for recognition and initial processing of extracellular signals. **Review Figure 6.20**

SELF-QUIZ

- Which statement about membrane phospholipids is *not* true?
 - They associate to form bilayers.
 - They have hydrophobic “tails.”
 - They have hydrophilic “heads.”
 - They give the membrane fluidity.
 - They flip-flop readily from one side of the membrane to the other.
- When a hormone molecule binds to a specific protein on the plasma membrane, the protein it binds to is called a
 - ligand.
 - clathrin.
 - receptor protein.
 - hydrophobic protein.
 - cell adhesion molecule.
- Which statement about membrane proteins is *not* true?
 - They all extend from one side of the membrane to the other.
 - Some serve as channels for ions to cross the membrane.
 - Many are free to migrate laterally within the membrane.
 - Their position in the membrane is determined by their tertiary structure.
 - Some play roles in photosynthesis.

4. Which statement about membrane carbohydrates is *not* true?
 - a. Some are bound to proteins.
 - b. Some are bound to lipids.
 - c. They are added to proteins in the Golgi apparatus.
 - d. They show little diversity.
 - e. They are important in recognition reactions at the cell surface.
 5. Which statement about animal cell junctions is *not* true?
 - a. Tight junctions are barriers to the passage of molecules between cells.
 - b. Desmosomes allow cells to adhere firmly to one another.
 - c. Gap junctions block communication between adjacent cells.
 - d. Connexons are made of protein.
 - e. The fibers associated with desmosomes are made of protein.
 6. You are studying how the protein transferrin enters cells. When you examine cells that have taken up transferrin, you find it inside clathrin-coated vesicles. Therefore, the most likely mechanism for uptake of transferrin is
 - a. facilitated diffusion.
 - b. an antiporter.
 - c. receptor-mediated endocytosis.
 - d. gap junctions.
 - e. ion channels.
7. Which statement about ion channels is *not* true?
 - a. They form pores in the membrane.
 - b. They are proteins.
 - c. All ions pass through the same type of channel.
 - d. Movement through them is from regions of high concentration to regions of low concentration.
 - e. Movement through them is by simple diffusion.
 8. Facilitated diffusion and active transport both
 - a. require ATP.
 - b. require the use of proteins as carriers or channels.
 - c. carry solutes in only one direction.
 - d. increase without limit as the concentration gradient increases.
 - e. depend on the solubility of the solute in lipids.
 9. Primary and secondary active transport both
 - a. generate ATP.
 - b. are based on passive movement of Na^+ ions.
 - c. include the passive movement of glucose molecules.
 - d. use ATP directly.
 - e. can move solutes against their concentration gradients.
 10. Which statement about osmosis is *not* true?
 - a. It obeys the laws of diffusion.
 - b. In animal tissues, water moves into cells if they are hypertonic to their environment.
 - c. Red blood cells must be kept in a plasma that is hypotonic to the cells.
 - d. Two cells with identical solute concentrations are isotonic to each other.
 - e. Solute concentration is the principal factor in osmosis.

FOR DISCUSSION

1. Muscle function requires calcium ions (Ca^{2+}) to be pumped into a subcellular compartment against a concentration gradient. What types of molecules are required for this to happen?
2. Section 27.5 describes the diatoms, which are protists that have complex glassy structures in their cell walls (see Figure 27.7B). These structures form within the Golgi apparatus. How do these structures reach the cell wall without having to pass through a membrane?
3. Organisms that live in fresh water are almost always hypertonic to their environment. In what way is this a serious problem? How do some organisms cope with this problem?
4. Contrast nonspecific endocytosis and receptor-mediated endocytosis.
5. The emergence of the phospholipid membrane was important to the origin of cells. Describe the properties of membranes that might have allowed cells to thrive in comparison with molecular aggregates without membranes.

ADDITIONAL INVESTIGATION

When a normal lung cell becomes a lung cancer cell, there are several important changes in plasma membrane properties. How would you investigate the following phenomena? (a) The cancer cell membrane is more fluid, with more rapid diffusion

in the plane of the membrane of both lipids and proteins. (b) The cancer cell has altered cell adhesion properties, binding to other tissues in addition to lung cells.

WORKING WITH DATA (GO TO yourBioPortal.com)

Aquaporin Increases Membrane Permeability to Water In this hands-on exercise based on Figure 6.13, you will investigate how Agre and colleagues used an egg cell to show that expression of aquaporin results in rapid water uptake when

the cell is placed in a hypotonic medium. Analyzing their experimental design and data, you will see how this model cell system and control experiments confirmed the important role of aquaporin as a water channel.

7

Cell Signaling and Communication

Love signals

Prairie voles (*Microtus ochrogaster*) are small rodents that live in temperate climates, where they dig tunnels in fields. When a male prairie vole encounters a female, mating often ensues. After mating (which can take as long as a day), the couple stays together, building a nest and raising their pups together. The two voles bond so tightly that they stay together for life. Contrast this behavior with that of the montane vole (*M. montanus*), which is closely related to the prairie vole and lives in the hills not far away. In this species, mating is quick, and afterwards the couple separates. The male looks for new mates and the female abandons her young soon after they are born.

The explanation for these dramatic behavioral differences lies in the brains of these two species. Neuroscientist Thomas Insel and his colleagues found that when prairie voles mate for all those hours, their brains release a

9-amino-acid peptide. In females, this peptide is oxytocin; in males, it is vasopressin. The peptide is circulated in the bloodstream and reaches all tissues in the body, but it binds to only a few cell types. These cells have surface proteins, called receptors, that specifically bind the peptide, like a key inserting into a lock.

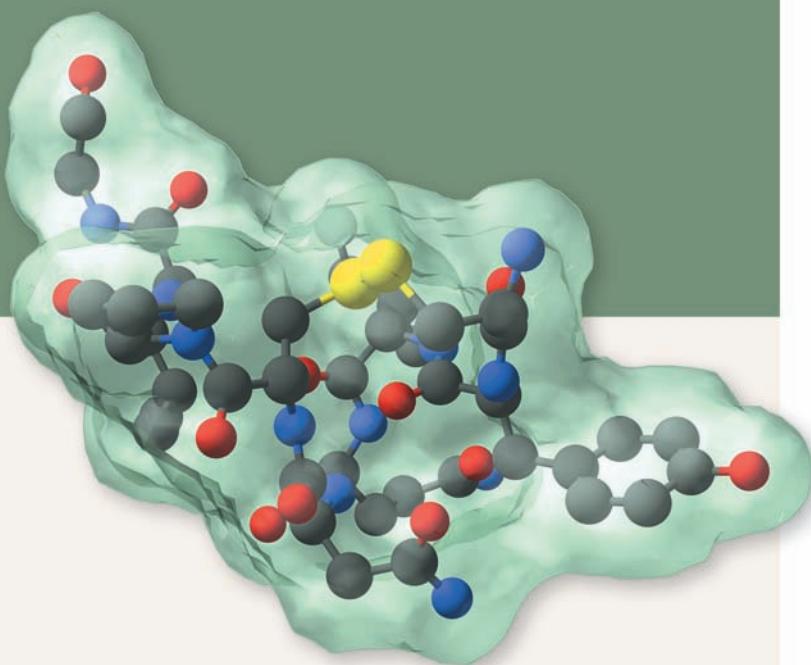
The interaction of peptide and receptor causes the receptor, which extends across the plasma membrane, to change shape. Within the cytoplasm, this change sets off a series of events called a signal transduction pathway. Such a pathway can cause many different cellular responses, but in this case, the notable changes are in behavior. The receptors for oxytocin and vasopressin in prairie voles are most concentrated in the regions of the brain that are responsible for behaviors such as bonding and caring for the young. In montane voles, there are far fewer receptors and

as a result, fewer postmating behaviors.

These cause-and-effect relationships between peptides, receptors, and behavior have been established through experiments. For example, a female prairie vole that is injected before mating with a molecule that blocks oxytocin does not bond with the male. Also, a female injected with oxytocin will bond with a male even without mating. Experiments with vasopressin in males give similar results. Furthermore, promiscuous vole males that were genetically manipulated to express prairie vole amounts of the vasopressin receptor grew up to behave more like prairie vole males. These experiments show that oxytocin and



Voles Prairie voles display extensive bonding behaviors after mating. These behaviors are mediated by peptides acting as intercellular signals.



Oxytocin This peptide with 9 amino acids acts as a signal for postmaturing behaviors.

vasopressin are signals that induce bonding and caring behaviors in voles. Could this also be true of humans?

Neuroeconomist Paul Zaks thinks so. He has done experiments with human volunteers, who were asked to “invest” funds with a stranger. A group of investors that was given a nasal spray containing oxytocin was more trusting of the stranger (and invested more funds) than a group that got an inert spray. So the oxytocin signaling pathway is important in human behavior too.

A cell’s response to any signal molecule takes place in three sequential steps. First, the signal binds to a receptor in the cell, often on the outside surface of the plasma membrane. Second, signal binding conveys a message to the cell. Third, the cell changes its activity in response to the signal. And in a multicellular organism, this leads to changes in that organism’s functioning.

IN THIS CHAPTER we first describe the types of signals that affect cells. These include chemicals produced by other cells and substances from outside the body, as well as physical and environmental factors such as light. Then we show how a signal affects only those cells that have the specific receptor to recognize that signal. Next, we describe the steps of signal transduction in which the receptor communicates to the cell that a signal has been received, thus causing a change in cell function.

CHAPTER OUTLINE

- 7.1 What Are Signals, and How Do Cells Respond to Them?
- 7.2 How Do Signal Receptors Initiate a Cellular Response?
- 7.3 How Is the Response to a Signal Transduced through the Cell?
- 7.4 How Do Cells Change in Response to Signals?
- 7.5 How Do Cells Communicate Directly?

7.1 What Are Signals, and How Do Cells Respond to Them?

Both prokaryotic and eukaryotic cells process information from their environments. This information can be in the form of a physical stimulus, such as the light reaching your eyes as you read this book, or chemicals that bathe a cell, such as lactose in a bacterial growth medium. It may come from outside the organism, such as the scent of a female moth seeking a mate in the dark, or from a neighboring cell within the organism, such as in the heart, where thousands of muscle cells contract in unison by transmitting signals to one another.

Of course, the mere presence of a signal does not mean that a cell will respond to it, just as you do not pay close attention to every image in your environment as you study. To respond to a signal, the cell must have a specific receptor that can detect it. This section provides examples of some types of cellular signals and one model of *signal transduction*. A **signal transduction pathway** is a sequence of molecular events and chemical reactions that lead to a cell’s response to a signal. After discussing signals in this section, we will consider their receptors in Section 7.2.

Cells receive signals from the physical environment and from other cells

The physical environment is full of signals. Our sense organs allow us to respond to light, odors and tastes (chemical signals), temperature, touch, and sound. Bacteria and protists can respond to minute chemical changes in their environments. Plants respond to light as a signal as well as an energy source. The amount and wavelengths of light reaching a plant’s surface differ from day to night and in direct sunlight versus shade. These variations act as signals that affect plant growth and reproduction. Some plants also respond to temperature: when the weather gets cold, they may respond either by becoming tolerant to cold or by accelerating flowering.

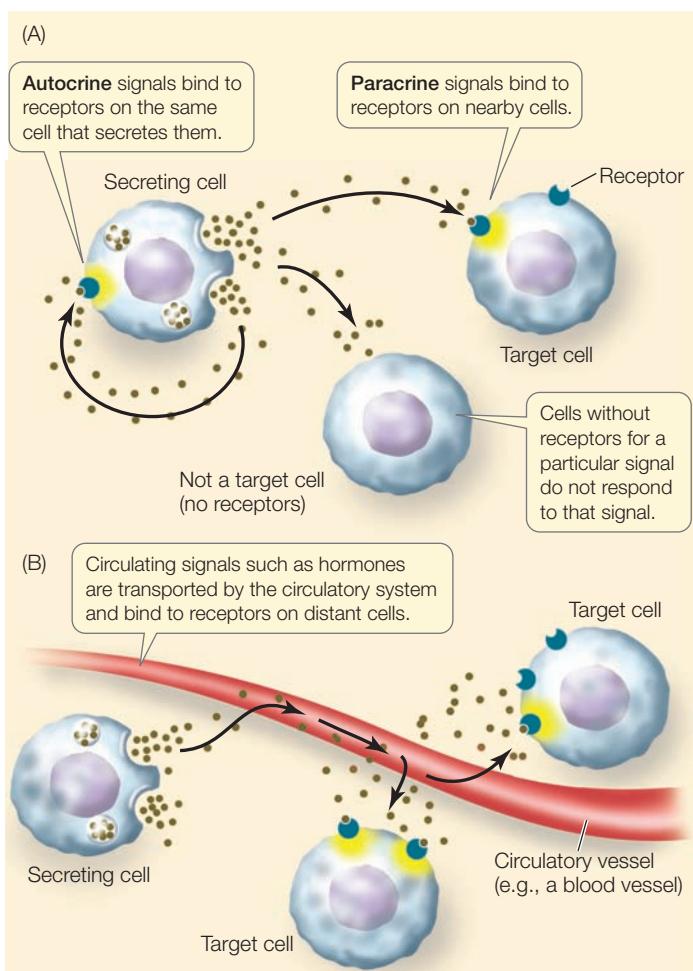
A cell deep inside a large multicellular organism is far away from the exterior environment. Such a cell’s environment consists of other cells and extracellular fluids. Cells receive their nutrients from, and pass their wastes into, extracellular fluids. Cells also receive signals—mostly chemical signals—from their extracellular fluid environment. Most of these chemical signals come from other cells, but they can also come from the environment via the digestive and respiratory systems. And cells can respond to changes in the extracellular concentrations of cer-

tain chemicals, such as CO_2 and H^+ , which are affected by the metabolic activities of other cells.

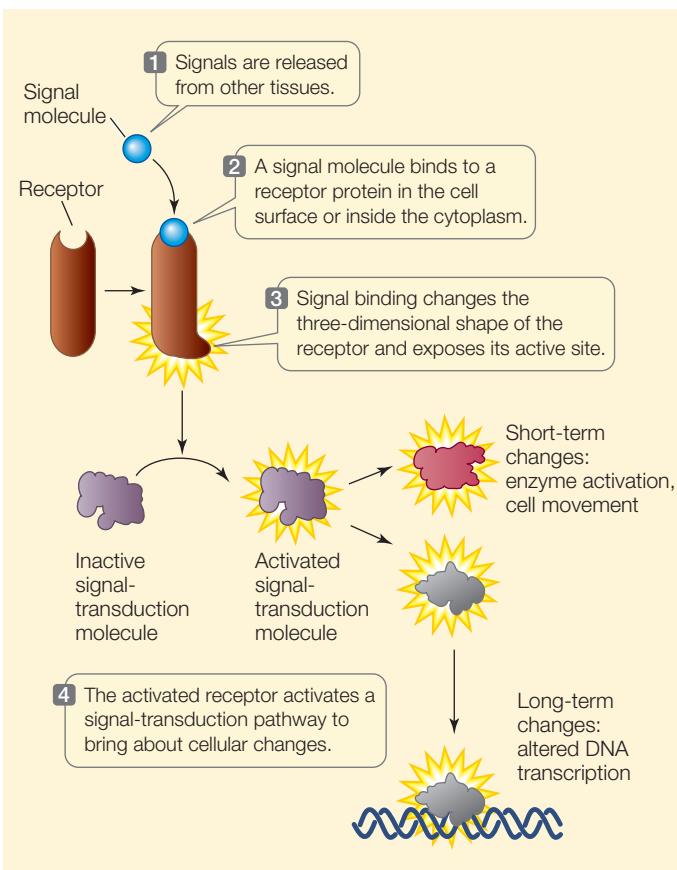
Inside a large multicellular organism, chemical signals made by the body itself reach a target cell by local diffusion or by circulation within the blood. These signals are usually in tiny concentrations (as low as 10^{-10} M) (see Chapter 2 for an explanation of *molar* concentrations). **Autocrine** signals diffuse to and affect the cells that make them; for example, part of the reason many tumor cells reproduce uncontrollably is because they self-stimulate cell division by making their own division signals. **Paracrine** signals diffuse to and affect nearby cells; an example is a neurotransmitter made by one nerve cell that diffuses to an adjacent cell and stimulates it. (**Figure 7.1A**). Signals to distant cells called hormones travel through the circulatory system (**Figure 7.1B**).

A signal transduction pathway involves a signal, a receptor, and responses

For the information from a signal to be transmitted to a cell, the target cell must be able to receive or sense the signal and respond to it, and the response must have some effect on the func-



7.1 Chemical Signaling Systems (A) A signal molecule can diffuse to act on the cell that produces it, or on a nearby cell. (B) Many signals act on distant cells and must be transported by the organism's circulatory system.



7.2 A Signal Transduction Pathway This general pathway is common to many cells and situations. The ultimate effects on the cell are either short-term or long-term molecular changes, or both.

tion of the cell. In a multicellular organism, all cells may receive chemical signals that are circulated in the blood, such as the peptides oxytocin and vasopressin that are released following mating in voles (see the opening of this chapter), but most body cells are not capable of responding to the signals. Only the cells with the necessary receptors can respond.

The kinds of responses vary greatly depending on the signal and the target cell. Just a few examples are: a skin cell initiating cell division to heal a wound; a cell moving to a new location in the embryo to form a tissue; a cell releasing enzymes to digest food; a plant cell loosening bonds that hold its cell wall polymers together so that it can expand; and a cell in the eye sending messages to the brain about the book you are reading. A signal transduction pathway involves a signal, a receptor, and a response (**Figure 7.2**).

Let's look at an example of such a pathway in the bacterium *Escherichia coli* (*E. coli*). Follow the features of this pathway in general (see Figure 7.2) and in particular (**Figure 7.3**).

SIGNAL As a prokaryotic cell, a bacterium is very sensitive to changes in its environment. One thing that can change is the total solute concentration (osmotic concentration—see Section 6.3) in the environment surrounding the cell. In the mammalian intestine where *E. coli* lives, the solute concentration around

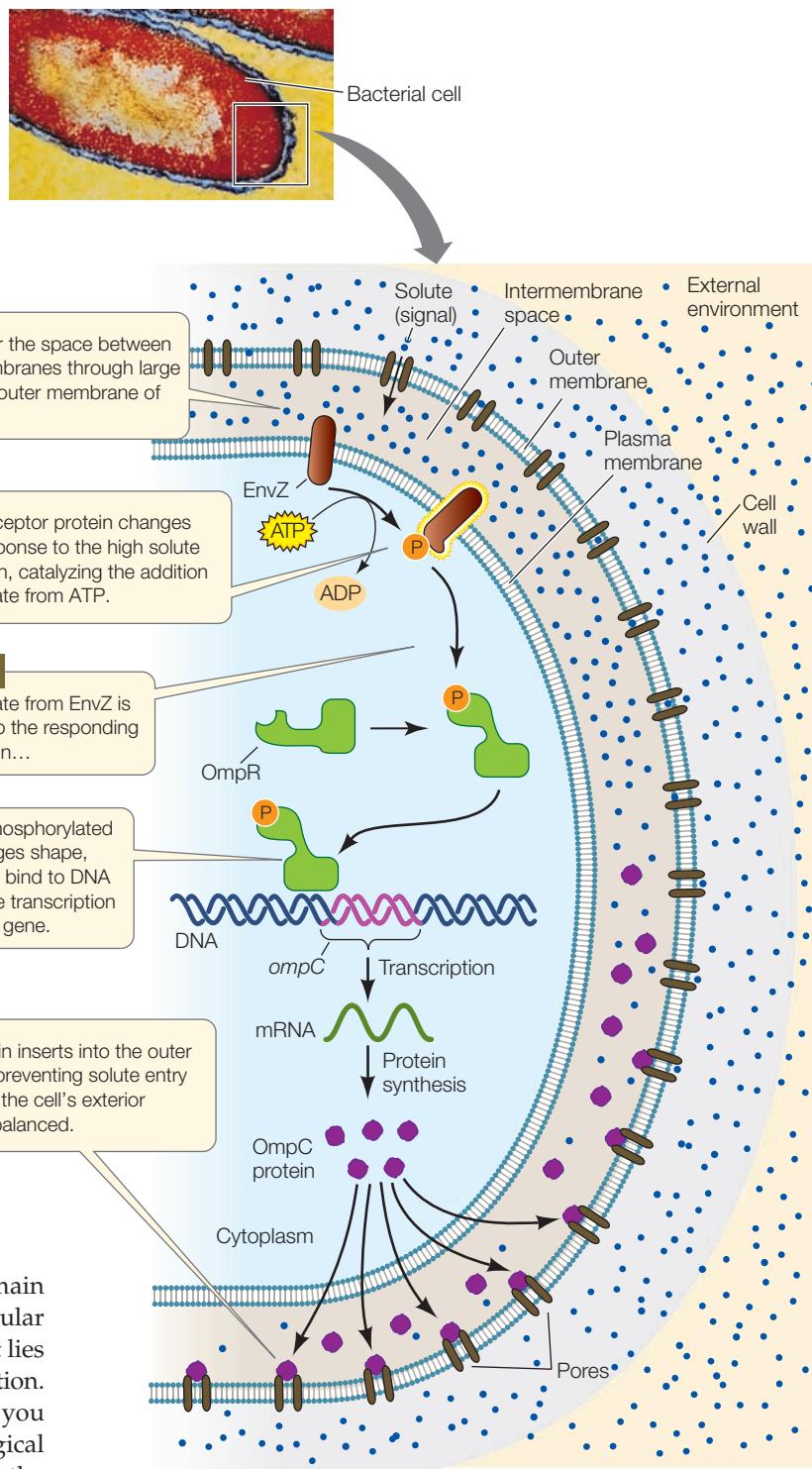
7.3 A Model Signal Transduction Pathway *E. coli* responds to the signal of an increase in solute concentration in its environment. The basic steps of such a signal transduction pathway occur in all living organisms.

yourBioPortal.com
GO TO Web Activity 7.1 • Signal Transduction

the bacterium often rises far above the solute concentration inside the cell. A fundamental characteristic of all living cells is that they maintain a constant internal environment, or homeostasis. To do this, the bacterium must perceive and quickly respond to this environmental signal (Figure 7.3, step 1). The cell does this by a signal transduction pathway involving two major components: a receptor and a responder.

RECEPTOR The *E. coli* receptor protein for changes in solute concentration is called EnvZ. EnvZ is a transmembrane protein that extends across the bacterium's plasma membrane into the space between the plasma membrane and the highly porous outer membrane, which forms a complex with the cell wall. When the solute concentration of the extracellular environment rises, so does the solute concentration in the space between the two membranes. This change in the aqueous solution causes the part of the receptor protein that sticks out into the intermembrane space to undergo a change in conformation (its three-dimensional shape).

The conformational change in the intermembrane domain (*a domain* is a sequence of amino acids folded into a particular shape) causes a conformational change in the domain that lies in the cytoplasm and initiates the events of signal transduction. The cytoplasmic domain of EnvZ can act as an *enzyme*. As you will see in more detail in Chapter 8, an enzyme is a biological catalyst that greatly speeds up a chemical reaction, and the active site is the region where the reaction actually takes place. The conformational change in EnvZ exposes an active site that was previously buried within the protein, so that EnvZ becomes a **protein kinase**—an enzyme that catalyzes the transfer of a phosphate group from ATP to another molecule. EnvZ transfers the phosphate group to one of its own histidine amino acids. In other words, EnvZ *phosphorylates* itself (Figure 7.3, step 2).



What does phosphorylation do to a protein? As discussed in Section 3.2, proteins can have both hydrophilic regions (which tend to interact with water on the outside of the protein macromolecule) and hydrophobic regions (which tend to interact with one another on the inside of the macromolecule). These regions are important in giving a protein its three-dimensional shape. Phosphate groups are charged, so an amino acid with such a group tends to be on the outside of the protein. Thus

phosphorylation leads to a change in the shape and function of a protein by changing its charge.

RESPONDER A **responder** is the second component of a signal transduction pathway. The charged phosphate group added to the histidine of the EnvZ protein causes its cytoplasmic domain to change its shape again. It now binds to a second protein, OmpR, and transfers the phosphate to it. In turn, this phosphorylation changes the shape of OmpR (**Figure 7.3, step 3**). The change in the responder is a key event in signaling, for three reasons:

- The signal on the outside of the cell has now been *transduced* to a protein that lies totally within the cell's cytoplasm.
- The altered responder can *do something*. In the case of the phosphorylated OmpR, that "something" is to bind to DNA to alter the expression of many genes; in particular, it increases the expression of the protein OmpC. This binding begins the final phase of the signaling pathway: the effect of the signal, which is an alteration in cell function.
- The signal has been *amplified*. Because a single enzyme can catalyze the conversion of many substrate molecules, one EnvZ molecule alters the structure of many OmpR molecules.

Phosphorylated OmpR has the correct three-dimensional structure to bind to the *ompC* DNA, resulting in an increase in the transcription of that gene. This results in the production of OmpC protein, which enables the cell to respond to the increase in osmotic concentration in its environment (**Figure 7.3, step 4**). The OmpC protein is inserted into the outer membrane of the cell, where it blocks pores and prevents solutes from entering the intermembrane space. As a result, the solute concentration in the intermembrane space is lowered, and homeostasis is restored. Thus the EnvZ-OmpR signal transduction pathway allows the *E. coli* cell to function just as if the external environment had a normal solute concentration.

Many of the elements that we have highlighted in this prokaryotic signal transduction pathway also exist in the signal transduction pathways of eukaryotic organisms. A typical eukaryotic signal transduction pathway has the following general steps:

- A receptor protein changes its conformation upon interaction with a signal. This receptor protein may or may not be in a membrane.
- A conformational change in the receptor protein activates its protein kinase activity, resulting in the transfer of a phosphate group from ATP to a target protein.
- This phosphorylation alters the function of a responder protein.
- The signal is amplified.
- A protein that binds to DNA is activated.
- The expression of one or more specific genes is turned on or off.
- Cell activity is altered.

7.1 RECAP

Cells are constantly exposed to molecular signals that can come from the external environment or from within the body of a multicellular organism. To respond to a signal, the cell must have a specific receptor that detects the signal and activates some cellular response.

- What are the differences between an autocrine signal, a paracrine signal, and a hormone? See p. 130 and **Figure 7.1**
- Describe the three components in a cell's response to a signal. See pp. 130–132 and **Figure 7.2**
- What are the elements of signal transduction that are described at the close of this section?

The general features of signal transduction pathways described in this section will recur in more detail throughout the chapter. First let's consider more closely the nature of the receptors that bind signal molecules.

7.2 How Do Signal Receptors Initiate a Cellular Response?

Any given cell in a multicellular organism is bombarded with many signals. However, it responds to only some of them, because no cell makes receptors for all signals. A receptor protein that binds to a chemical signal does so very specifically, in much the same way that a membrane transport protein binds to the substance it transports. This *specificity* of binding ensures that only those cells that make a specific receptor will respond to a given signal.

Receptors have specific binding sites for their signals

A specific chemical signal molecule fits into a three-dimensional site on its protein receptor (**Figure 7.4A**). A molecule that binds to a receptor site on another molecule in this way is called a **ligand**. Binding of the signaling ligand causes the receptor protein to change its three-dimensional shape, and that conformational change initiates a cellular response. The ligand does not contribute further to this response. In fact, the ligand is usually not metabolized into a useful product; its role is purely to "knock on the door." (This is in sharp contrast to the enzyme–substrate interaction, which is described in Chapter 8. The whole purpose of that interaction is to change the substrate into a useful product.)

Receptors bind to their ligands according to chemistry's *law of mass action*:



This means that the binding is reversible, although for most ligand–receptor complexes, the equilibrium point is far to the right—that is, binding is favored. Reversibility is important, however, because if the ligand were never released, the receptor would be continuously stimulated.

7.4 A Signal and Its Receptor (A) The adenosine 2A receptor occurs in the human brain, where it is involved in inhibiting arousal. (B) Adenosine is the normal ligand for the receptor. Caffeine has a similar structure to that of adenosine and can act as an antagonist that binds the receptor and prevents its normal functioning.

An inhibitor (or *antagonist*) can also bind to a receptor protein, instead of the normal ligand. There are both natural and artificial antagonists of receptor binding. For example, many substances that alter human behavior bind to specific receptors in the brain, and prevent the binding of the receptors' specific ligands. An example is caffeine, which is probably the world's most widely consumed stimulant. In the brain, the nucleoside adenosine acts as a ligand that binds to a receptor on nerve cells, initiating a signal transduction pathway that reduces brain activity, especially arousal. Because caffeine has a similar molecular structure to that of adenosine, it also binds to the adenosine receptor (**Figure 7.4B**). But in this case binding does not initiate a signal transduction pathway. Rather, it "ties up" the receptor, preventing adenosine binding and thereby allowing nerve cell activity and arousal.

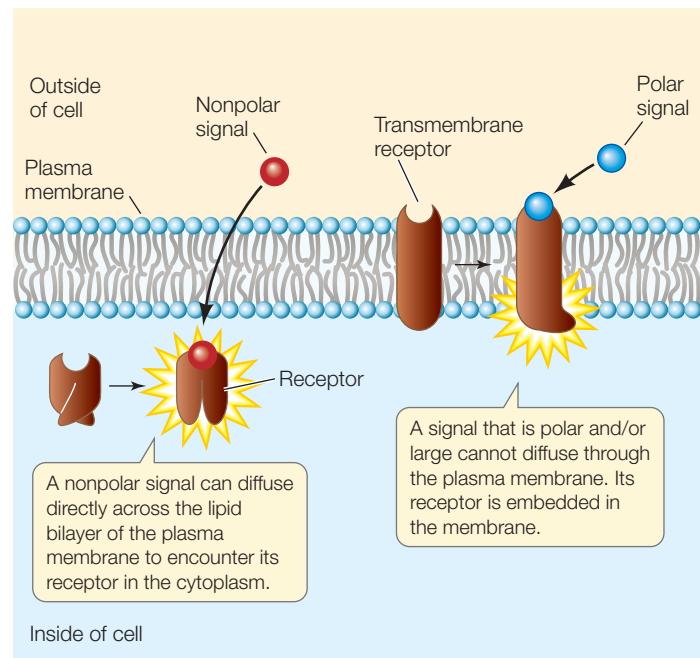
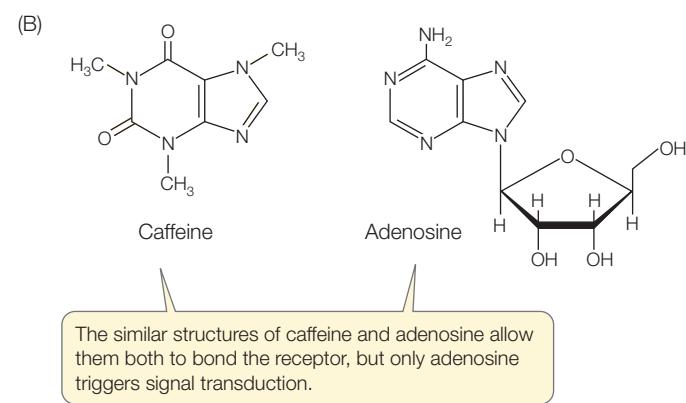
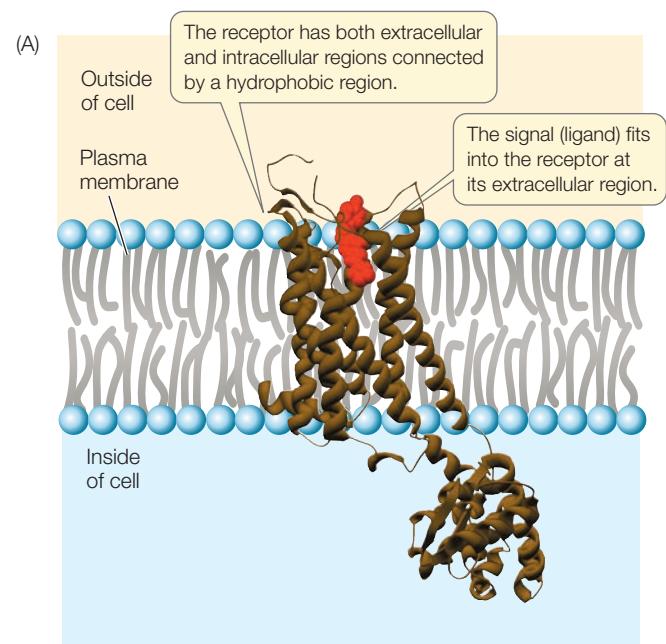
Receptors can be classified by location and function

The chemistry of ligand signals is quite variable, but they can be divided into two groups, based on whether or not they can diffuse through membranes. Correspondingly, a receptor can be classified by its location in the cell, which largely depends on the nature of its ligand (**Figure 7.5**):

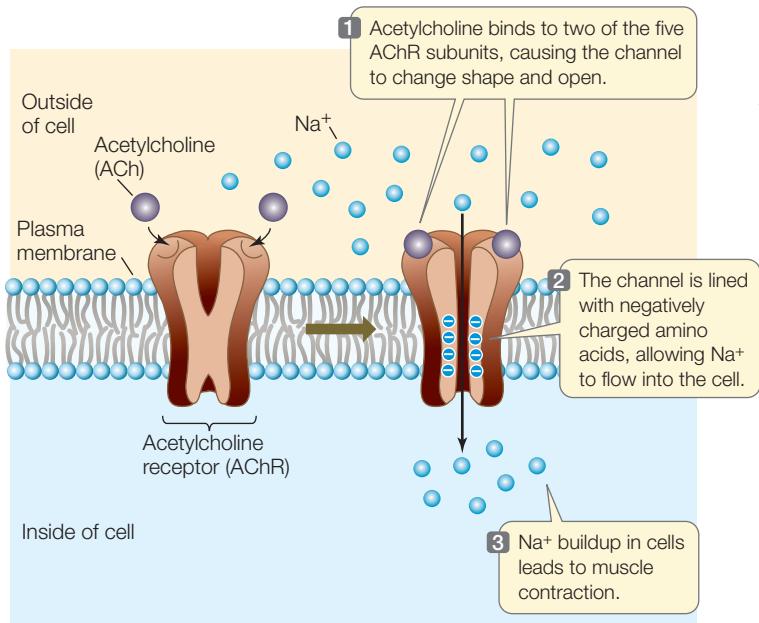
- **Cytoplasmic receptors:** Small or nonpolar ligands can diffuse across the nonpolar phospholipid bilayer of the plasma membrane and enter the cell. Estrogen, for example, is a lipid-soluble steroid hormone that can easily diffuse across the plasma membrane; it binds to a receptor in the cytoplasm.
- **Membrane receptors:** Large or polar ligands cannot cross the lipid bilayer. Insulin, for example, is a protein hormone that cannot diffuse through the plasma membrane; instead, it binds to a transmembrane receptor with an extracellular binding domain.

In complex eukaryotes such as mammals and higher plants, there are three well-studied categories of plasma membrane receptors that are grouped according to their functions: ion channels, protein kinase receptors, and G protein-linked receptors.

ION CHANNEL RECEPTORS As described in Section 6.3, the plasma membranes of many types of cells contain gated **ion channels** for ions such as Na^+ , K^+ , Ca^{2+} , or Cl^- to enter or leave the cell (see Figure 6.11). The gate-opening mechanism is an alteration in the three-dimensional shape of the channel protein upon ligand binding; thus these proteins function as receptors. Each type of ion channel has its own signal, and these include sensory stimuli such as light, sound, and electric charge

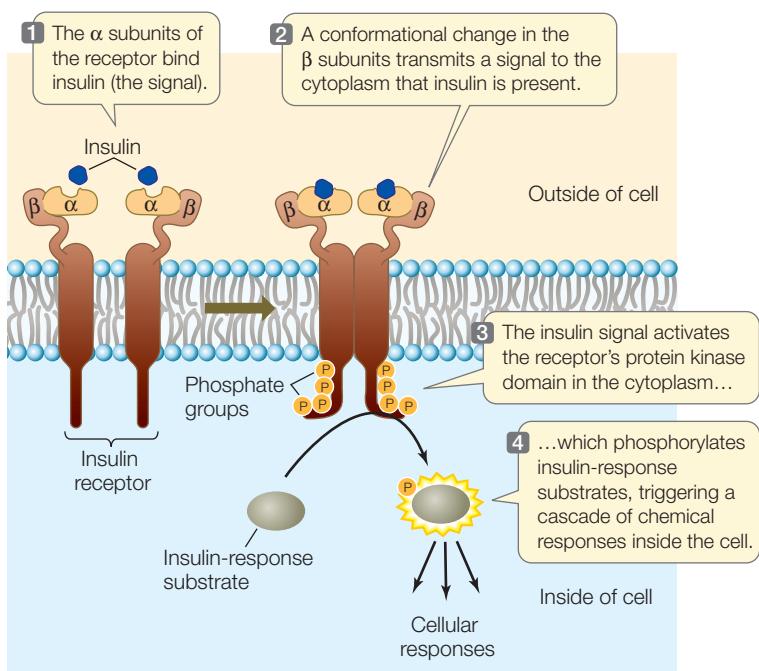


7.5 Two Locations for Receptors Receptors can be located in the cytoplasm or in the plasma membrane of the cell.



differences across the plasma membrane, as well as chemical ligands such as hormones and neurotransmitters.

The *acetylcholine receptor*, which is located in the plasma membrane of skeletal muscle cells, is an example of a gated ion channel. This receptor protein is a sodium channel that binds the ligand acetylcholine, which is a neurotransmitter—a chemical signal released from neurons (nerve cells) (Figure 7.6). When two molecules of acetylcholine bind to the receptor, it opens for about a thousandth of a second. That is enough time for Na^+ , which is more concentrated outside the cell than inside, to rush into the cell, moving in response to both concentration and electrical potential gradients. The change in Na^+ concentration in the cell initiates a series of events that result in muscle contraction.



7.6 A Gated Ion Channel The acetylcholine receptor (AChR) is a ligand-gated ion channel for sodium ions. It is made up of five polypeptide subunits. When acetylcholine molecules (ACh) bind to two of the subunits, the gate opens and Na^+ flows into the cell. This channel helps regulate membrane polarity (see Chapter 6).

PROTEIN KINASE RECEPTORS Like the EnvZ receptor of *E. coli*, some eukaryotic receptor proteins become protein kinases when they are activated. They catalyze the phosphorylation of themselves and/or other proteins, thus changing their shapes and therefore their functions.

The receptor for insulin is an example of a protein kinase receptor. Insulin is a protein hormone made by the mammalian pancreas. Its receptor has two copies each of two different polypeptide subunits (Figure 7.7). When insulin binds to the receptor, the receptor becomes activated and able to phosphorylate itself and certain cytoplasmic proteins that are appropriately called *insulin response substrates*. These proteins then initiate many cellular responses, including the insertion of glucose transporters (see Figure 6.14) into the plasma membrane.

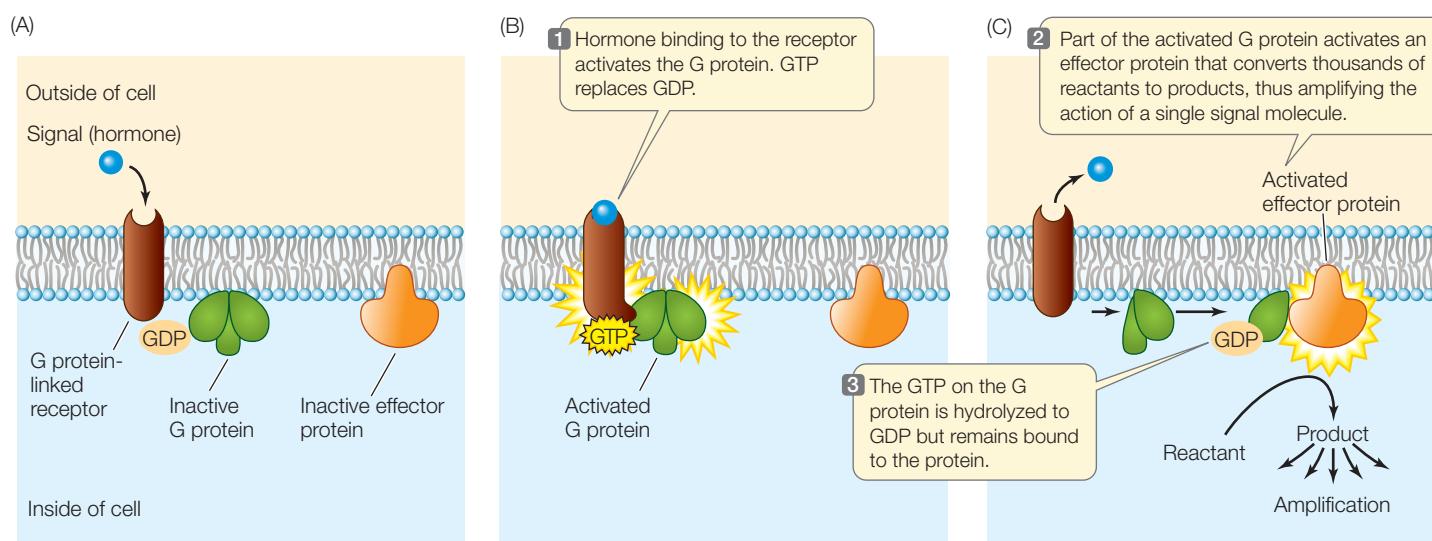
G PROTEIN-LINKED RECEPTORS A third category of eukaryotic plasma membrane receptors is the G protein-linked receptors, also referred to as the seven transmembrane domain receptors. This descriptive name identifies a fascinating group of receptors, each of which is composed of a single protein with seven transmembrane domains. These seven domains pass through the phospholipid bilayer and are separated by short loops that extend either outside or inside the cell. Ligand binding on the extracellular side of the receptor changes the shape of its cytoplasmic region, exposing a site that binds to a mobile membrane protein called a **G protein**. The G protein is partially inserted into the lipid bilayer and partially exposed on the cytoplasmic surface of the membrane.

Many G proteins have three polypeptide subunits and can bind three different molecules (Figure 7.8A):

- The receptor
- GDP and GTP (guanosine diphosphate and triphosphate, respectively; these are nucleoside phosphates like ADP and ATP)
- An effector protein

When the G protein binds to an activated receptor protein, GDP is exchanged for GTP (Figure 7.8B). At the same time, the ligand is usually released from the extracellular side of the receptor. GTP binding causes a conformational change in the G protein. The GTP-bound subunit then separates from the rest of the protein, diffusing in the plane of the phospholipid bilayer until it encounters an **effector protein** to which it can bind. An effector protein is just what its name implies: it causes an effect in the cell. The binding of the GTP-bearing G protein

7.7 A Protein Kinase Receptor The mammalian hormone insulin binds to a receptor on the outside surface of the cell and initiates a response.



7.8 A G Protein-Linked Receptor The G protein is an intermediary between the receptor and an effector.

yourBioPortal.com

GO TO Animated Tutorial 7.1 • Signal Transduction Pathway

subunit activates the effector—which may be an enzyme or an ion channel—thereby causing changes in cell function (**Figure 7.8C**).

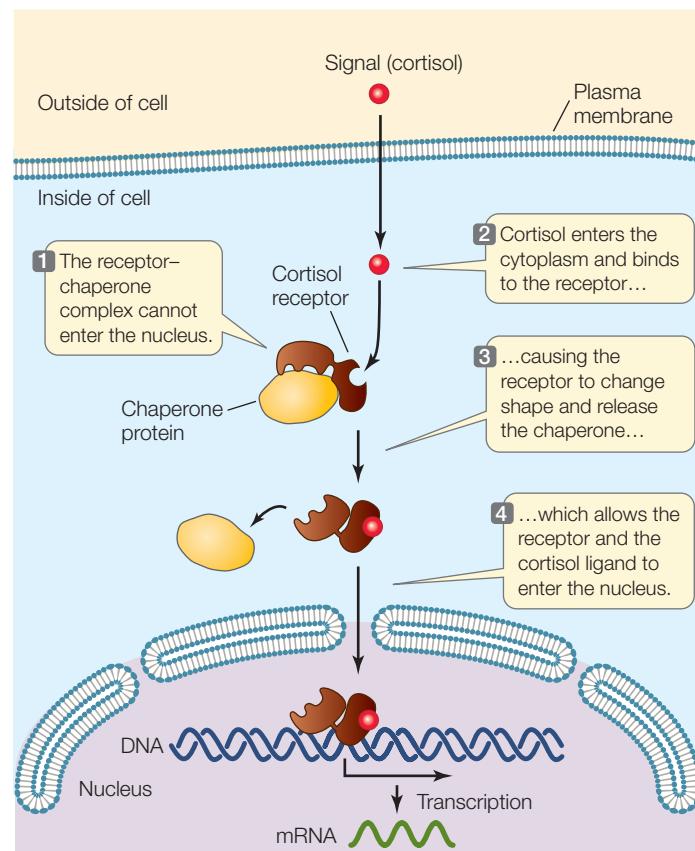
After activation of the effector protein, the GTP on the G protein is hydrolyzed to GDP. The now inactive G protein subunit separates from the effector protein and diffuses in the membrane to collide with and bind to the other two G protein subunits. When the three components of the G protein are reassembled, the protein is capable of binding again to an activated receptor. After binding, the activated receptor exchanges the GDP on the G protein for a GTP, and the cycle begins again.

There are variations in all three G protein subunits, giving different G protein complexes different functions. A G protein can either activate or inhibit an effector protein. An example in humans of an *activating* response involves the receptor for epinephrine (adrenaline), which is a hormone made by the adrenal gland in response to stress or heavy exercise. In heart muscle, this hormone binds to its G protein-linked receptor, activating a G protein. The GTP-bound subunit then activates a membrane-bound enzyme to produce a small molecule, cyclic adenosine monophosphate (cAMP). This molecule, in turn, has many effects on the cell (as we will see below), including the mobilization of glucose for energy and muscle contraction.

G protein-mediated *inhibition* occurs when the same hormone, epinephrine, binds to its receptor in the smooth muscle cells surrounding blood vessels lining the digestive tract. Again, the epinephrine-bound receptor changes its shape and activates a G protein, and the GTP-bound subunit binds to a target enzyme. But in this case, the enzyme is inhibited instead of being activated. As a result, the muscles relax and the blood vessel diameter increases, allowing more nutrients to be carried away from the digestive system to the rest of the body. Thus the same signal and signaling mechanism can have different consequences in different cells, depending on the presence of specific receptor and effector molecules.

CYTOPLASMIC RECEPTORS **Cytoplasmic receptors** are located inside the cell and bind to signals that can diffuse across the

plasma membrane. Binding to the signaling ligand causes the receptor to change its shape so that it can enter the cell nucleus, where it affects expression of specific genes. But this general view is somewhat simplified. The receptor for the steroid hormone cortisol, for example, is normally bound to a chaperone protein, which blocks it from entering the nucleus. Binding of the hormone causes the receptor to change its shape so that the chaperone is released (**Figure 7.9**). This release allows the



7.9 A Cytoplasmic Receptor The receptor for cortisol is bound to a chaperone protein. Binding of the signal to the receptor releases the chaperone and allows the ligand-receptor complex to enter the cell's nucleus, where it binds to DNA. Changes in DNA transcription are long-term in comparison to the more immediate changes in enzyme activity observed in other pathways (see Figure 7.20).

receptor to fold into an appropriate conformation for entering the nucleus and initiating DNA transcription.

7.2 RECAP

Receptors are proteins that bind, or are changed by, specific signals or ligands; the changed receptor initiates a response in the cell. These receptors may be at the plasma membrane or inside the cell.

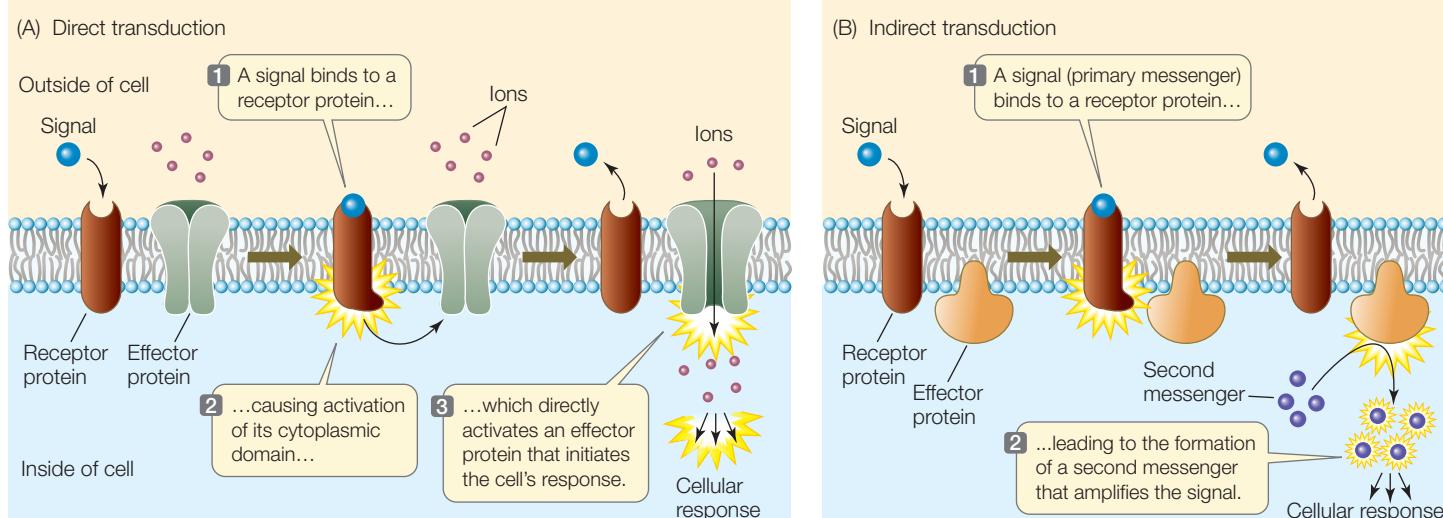
- What are the nature and importance of specificity in the binding of receptors to their particular ligands? **See pp. 132–133**
- What are three important categories of plasma membrane receptors seen in complex eukaryotes? **See pp. 133–134 and Figures 7.6, 7.7, and 7.8**

Now that we have discussed signals and receptors, let's examine the characteristics of the molecules (*transducers*) that mediate between the receptor and the cellular response.

7.3 How Is the Response to a Signal Transduced through the Cell?

As we have just seen with epinephrine, the same signal may produce different responses in different tissues. These different responses to the same signal–receptor complex are mediated by the components of different signal transduction pathways. Signal transduction may be either direct or indirect:

- **Direct transduction** is a function of the receptor itself and occurs at the plasma membrane. The interaction between the signal (primary messenger) and receptor results in the cellular response. (Figure 7.10A).
- In **indirect transduction**, which is more common, another molecule termed a **second messenger** diffuses into the cytoplasm and mediates additional steps in the signal transduction pathway (Figure 7.10B).



In both cases, the signal can initiate a *cascade* of events, in which proteins interact with other proteins until the final responses are achieved. Through such a cascade, an initial signal can be both amplified and distributed to cause several different responses in the target cell.

A protein kinase cascade amplifies a response to ligand binding

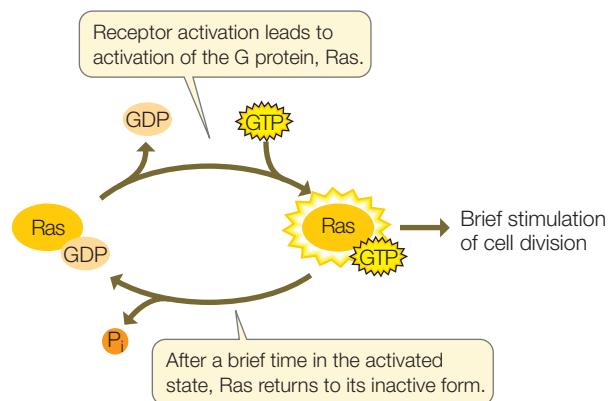
We have seen that when a signal binds to a protein kinase receptor, the receptor's conformation changes, exposing a protein kinase active site on the receptor's cytoplasmic domain. The protein kinase then catalyzes the phosphorylation of target proteins. This process is an example of direct signal transduction, because the amplifying enzyme is the receptor itself. Protein kinase receptors are important in binding signals called growth factors that stimulate cell division in both plants and animals.

A complete signal transduction pathway that occurs after a protein kinase receptor binds a growth factor was discovered in studies on a cell that went wrong. Many human bladder cancers contain an abnormal form of a protein called Ras (so named because a similar protein was previously isolated from a *rat* sarcoma tumor). Investigations of these bladder cancers showed that Ras was a G protein, and the abnormal form was always active because it was permanently bound to GTP, and thus caused continuous cell division (Figure 7.11). If this abnormal form of Ras was inhibited, the cells stopped dividing. This discovery has led to a major effort to develop specific Ras inhibitors for cancer treatment.

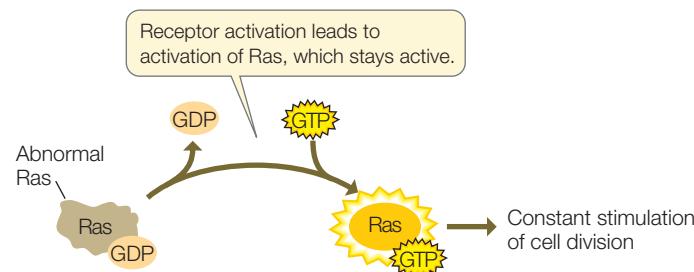
Other cancers have abnormalities in different aspects of signal transduction. Biologists have compared the defects in these cells with the normal signaling process in non-cancerous cells, and thus worked out the entire signaling pathway. It is an ex-

7.10 Direct and Indirect Signal Transduction (A) All the events of direct transduction occur at or near the receptor (in this case, at the plasma membrane). (B) In indirect transduction, a second messenger mediates the events inside the cell. The signal is considered to be the first messenger.

(A) Normal cell



(B) Cancer cell



7.11 Signal Transduction and Cancer (A) Ras is a G protein that regulates cell division. (B) In some tumors, the Ras protein is permanently active, resulting in uncontrolled cell division.

ample of a more general phenomenon, called a **protein kinase cascade**, where one protein kinase activates the next, and so on (Figure 7.12). Such cascades are key to the external regulation of many cellular activities. Indeed, the eukaryotic genome codes for hundreds, even thousands, of such kinases.

Protein kinase cascades are useful signal transducers for four reasons:

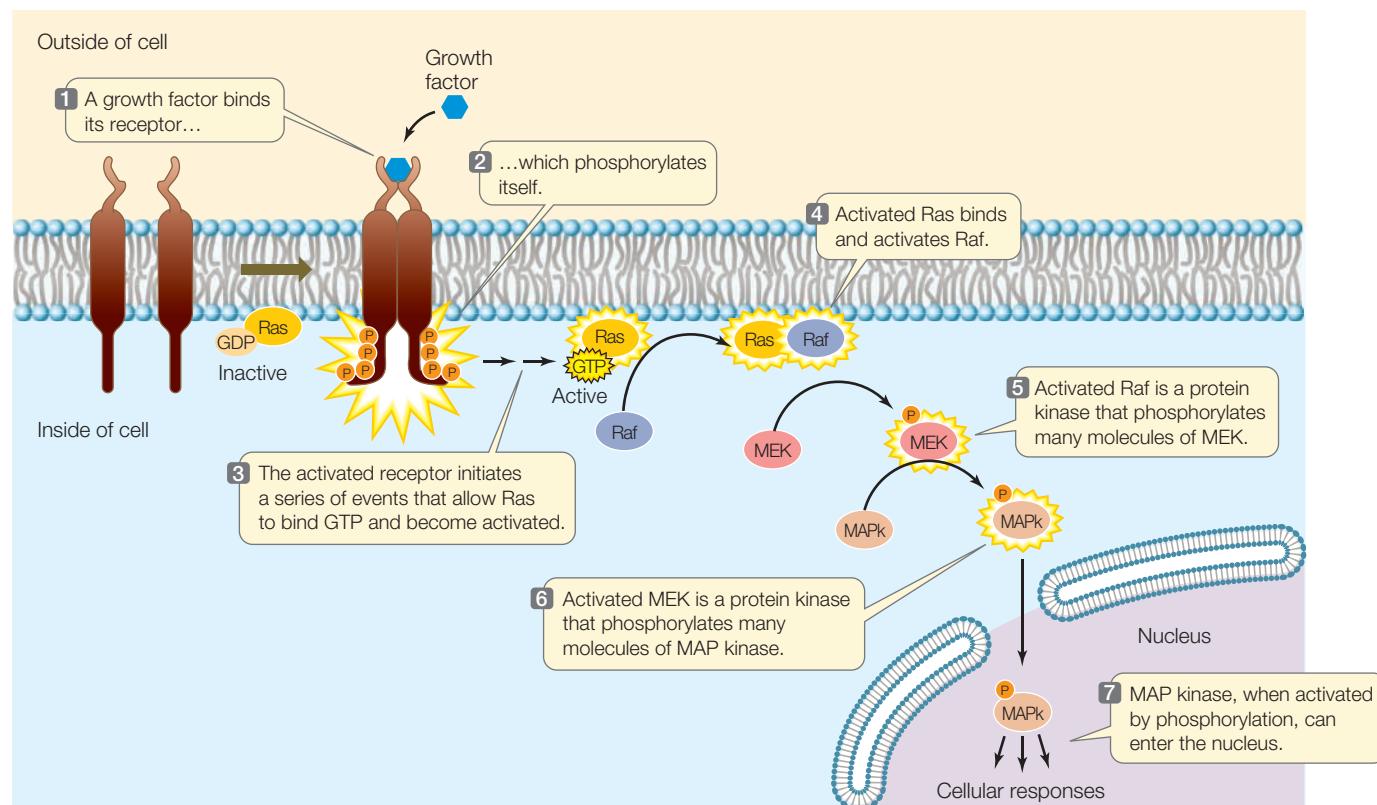
- At each step in the cascade of events, the signal is *amplified*, because each newly activated protein kinase is an enzyme that can catalyze the phosphorylation of many target proteins.
- The information from a signal that originally arrived at the plasma membrane is *communicated* to the nucleus.
- The multitude of steps provides some *specificity* to the process.
- Different target proteins at each step in the cascade can provide *variation* in the response.

yourBioPortal.com

GO TO Animated Tutorial 7.2 • Signal Transduction and Cancer

Second messengers can stimulate protein kinase cascades

As we have just seen, protein kinase receptors initiate protein kinase cascades right at the plasma membrane. However, the stimulation of events in the cell is more often indirect. In a series of clever experiments, Earl Sutherland and his colleagues at Case Western Reserve University discovered that a small water-solu-



7.12 A Protein Kinase Cascade In a protein kinase cascade, a series of proteins are sequentially activated.

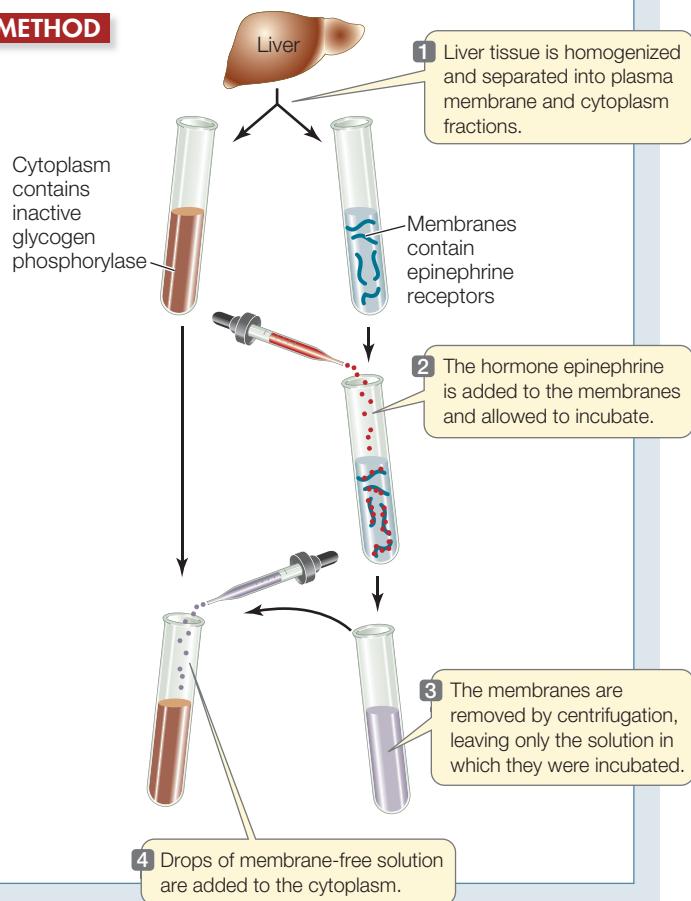
INVESTIGATING LIFE

7.13 The Discovery of a Second Messenger

Glycogen phosphorylase is activated in liver cells after epinephrine binds to a membrane receptor. Sutherland and his colleagues observed that this activation could occur *in vivo* only if fragments of the plasma membrane were present. They designed experiments to show that a second messenger caused the activation of glycogen phosphorylase.

HYPOTHESIS A second messenger mediates between receptor activation at the plasma membrane and enzyme activation in the cytoplasm.

METHOD



RESULTS

Active glycogen phosphorylase is present in the cytoplasm.

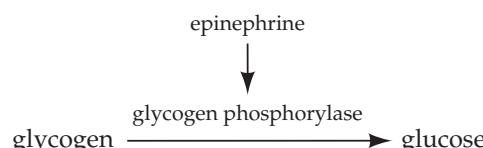
CONCLUSION

A soluble second messenger, produced by hormone-activated membranes, is present in the solution and activates enzymes in the cytoplasm.

FURTHER INVESTIGATION: The soluble molecule produced in this experiment was later identified as cAMP. How would you show that cAMP, and not ATP, is the second messenger in this system?

Go to [yourBioPortal.com](#) for original citations, discussions, and relevant links for all INVESTIGATING LIFE figures.

ble chemical messenger mediates the cytoplasmic events initiated by a plasma membrane receptor. These researchers were investigating the activation of the liver enzyme glycogen phosphorylase by the hormone epinephrine. The enzyme is released when an animal faces life-threatening conditions and needs energy fast for the fight-or-flight response. Glycogen phosphorylase catalyzes the breakdown of glycogen stored in the liver so that the resulting glucose molecules can be released to the blood. The enzyme is present in the liver cell cytoplasm, but is inactive except in the presence of epinephrine.

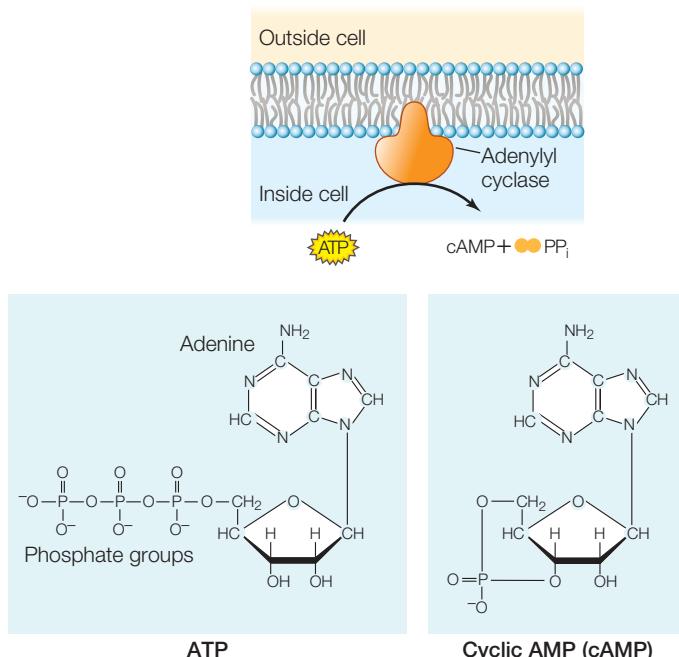


The researchers found that epinephrine could activate glycogen phosphorylase in liver cells that had been broken open, but only if the entire cell contents, including plasma membrane fragments, were present. Under these circumstances epinephrine bound to the plasma membranes, but the active phosphorylase was present in the solution. The researchers hypothesized that there must be a second “messenger” that transmits the signal of epinephrine (the “first messenger,” which binds to a receptor at the plasma membrane) to the phosphorylase (in the cytoplasm). To investigate the production of this messenger, they separated plasma membrane fragments from the cytoplasms of broken liver cells and followed the sequence of steps described in **Figure 7.13**. This experiment confirmed their hypothesis that hormone binding to the membrane receptor causes the production of a small, water-soluble molecule that diffuses into the cytoplasm and activates the enzyme. Later, this second messenger was identified as **cyclic AMP (cAMP)**. (We will describe the signal transduction pathway leading to the fight-or-flight response in more detail in Section 7.4.) Second messengers do not have enzymatic activity; rather, they act to regulate target enzymes (see Chapter 8).

A second messenger is a small molecule that mediates later steps in a signal transduction pathway after the first messenger—the signal or ligand—binds to its receptor. In contrast to the specificity of receptor binding, second messengers allow a cell to respond to a single event at the plasma membrane with *many events inside the cell*. Thus, second messengers serve to amplify the signal—for example, binding of a single epinephrine molecule leads to the production of many molecules of cAMP, which then activate many enzyme targets by binding to them noncovalently. In the case of epinephrine and the liver cell, glycogen phosphorylase is just one of several enzymes that are activated.

Cyclic AMP is a second messenger in a wide variety of signal transduction pathways. An effector protein, adenylyl cyclase, catalyzes the formation of cAMP from ATP. Adenylyl cyclase is located on the cytoplasmic surface of the plasma membrane of target cells (**Figure 7.14**). Usually a G protein activates the enzyme after it has itself been activated by a receptor.

Cyclic AMP has two major kinds of targets—ion channels and protein kinases. In many sensory cells, cAMP binds to ion channels and thus opens them. Cyclic AMP may also bind to a



7.14 The Formation of Cyclic AMP The formation of cAMP from ATP is catalyzed by adenyl cyclase, an enzyme that is activated by G proteins.

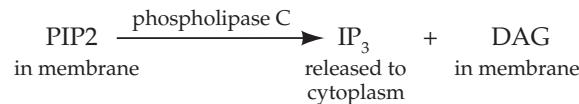
protein kinase in the cytoplasm, activating its catalytic function. A protein kinase cascade (see Figure 7.12) ensues, leading to the final effects in the cell.

Second messengers can be derived from lipids

In addition to their role as structural components of the plasma membrane, phospholipids are also involved in signal transduction. When certain phospholipids are hydrolyzed into their component parts by enzymes called **phospholipases**, second messengers are formed.

The best-studied examples of lipid-derived second messengers come from the hydrolysis of the phospholipid **phosphatidyl inositol-bisphosphate (PIP2)**. Like all phospholipids, PIP2 has a hydrophobic portion embedded in the plasma membrane: two fatty acid tails attached to a molecule of glycerol, which together form **diacylglycerol**, or **DAG**. The hydrophilic portion of PIP2 is **inositol trisphosphate**, or **IP₃**, which projects into the cytoplasm.

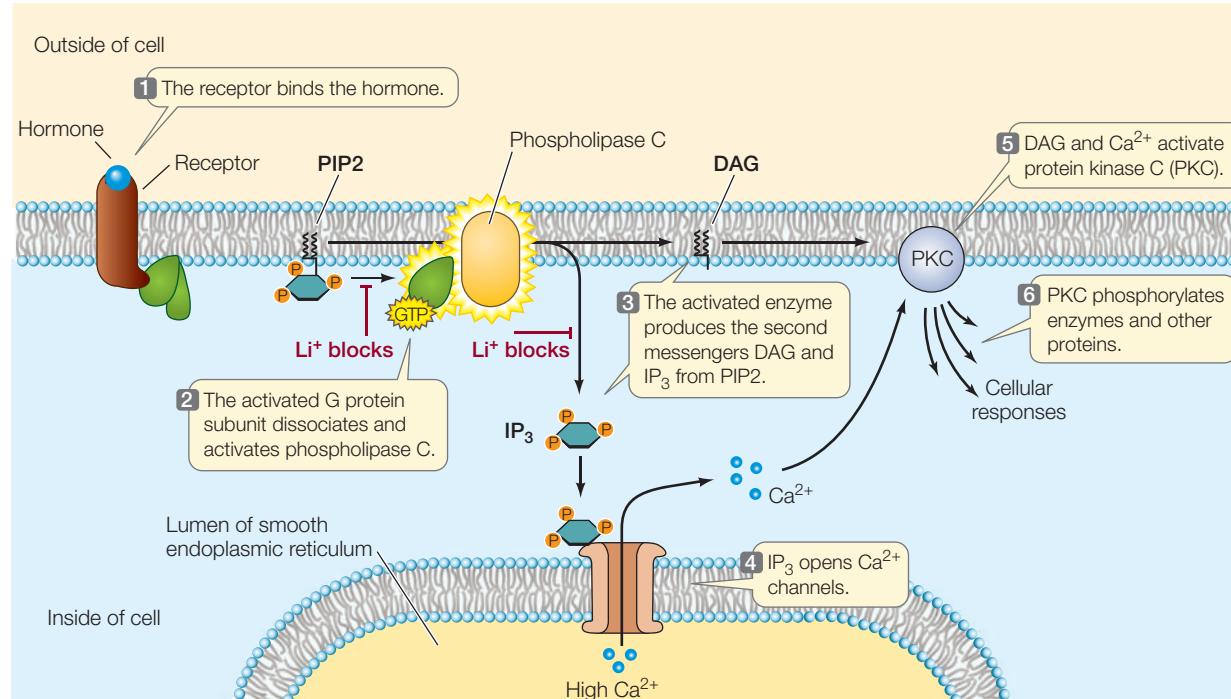
As with cAMP, the receptors involved in this second-messenger system are often G protein-linked receptors. A G protein subunit is activated by the receptor, then diffuses within the plasma membrane and activates phospholipase C, an enzyme that is also located in the membrane. This enzyme cleaves off the IP₃ from PIP2, leaving the diacylglycerol (DAG) in the phospholipid bilayer:



IP₃ and DAG, both second messengers, have different modes of action that build on each other, activating protein kinase C (PKC) (Figure 7.15). PKC refers to a family of protein kinases that can phosphorylate a wide variety of target proteins, leading to a multiplicity of cellular responses that vary depending on the tissue or cell type.

The IP₃/DAG pathway is apparently a target for the ion lithium (Li⁺), which was used for many years as a psychoactive drug to treat bipolar (manic-depressive) disorder. This serious illness occurs in about 1 in every 100 people. In these patients, an overactive IP₃/DAG signal transduction pathway in the

7.15 The IP₃/DAG Second-Messenger System Phospholipase C hydrolyzes the phospholipid PIP2 into its components, IP₃ and DAG, both of which are second messengers. Lithium ions (Li⁺) block this pathway and are used to treat bipolar disorder (red type).



brain leads to excessive brain activity in certain regions. Lithium “tones down” this pathway in two ways, as indicated by the red notations in Figure 7.15. It inhibits G protein activation of phospholipase C, and also inhibits the synthesis of IP₃. The overall result is that brain activity returns to normal.

Calcium ions are involved in many signal transduction pathways

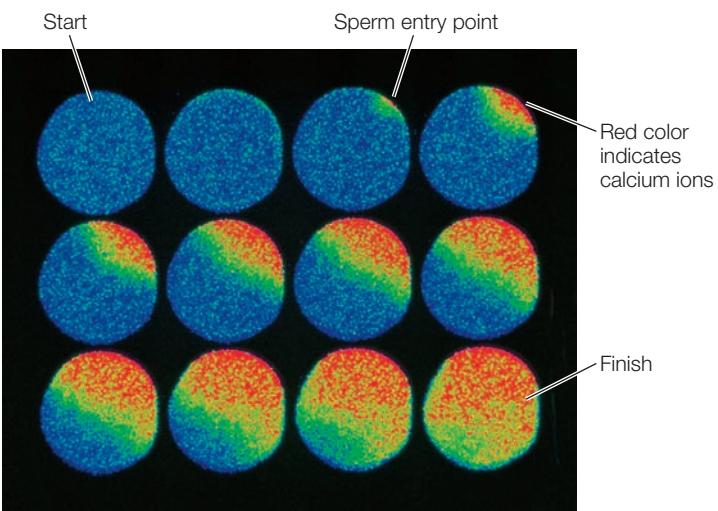
Calcium ions (Ca²⁺) are scarce inside most cells, which have cytosolic Ca²⁺ concentrations of only about 0.1 mM. Ca²⁺ concentrations outside cells and within the endoplasmic reticulum are usually much higher. Active transport proteins in the plasma and ER membranes maintain this concentration difference by pumping Ca²⁺ out of the cytosol. In contrast to cAMP and the lipid-derived second messengers, Ca²⁺ cannot be made in order to increase the intracellular Ca²⁺ concentration. Instead, Ca²⁺ ion levels are regulated via the opening and closing of ion channels, and the action of membrane pumps.

There are many signals that can cause calcium channels to open, including IP₃ (see Figure 7.15). The entry of a sperm into an egg is a very important signal that causes a massive opening of calcium channels, resulting in numerous and dramatic changes that prepare the now fertilized egg for cell divisions and development (Figure 7.16). Whatever the initial signal that causes the calcium channels to open, their opening results in a dramatic increase in cytosolic Ca²⁺ concentration, which can increase up to one hundredfold within a fraction of a second. As we saw earlier, this increase activates protein kinase C. In addition, Ca²⁺ controls other ion channels and stimulates secretion by exocytosis in many cell types.

Nitric oxide can act in signal transduction

Most signaling molecules and second messengers are solutes that remain dissolved in either the aqueous or hydrophobic components of cells. It was a great surprise to find that a gas could also be active in signal transduction. Pharmacologist Robert Furchtgott, at the State University of New York in Brooklyn, was investigating the mechanisms that cause the smooth muscles lining blood vessels in mammals to relax, thus allowing more blood to flow to certain organs. The neurotransmitter acetylcholine (see Section 7.2) appeared to stimulate the IP₃/DAG signal transduction pathway to produce an influx of Ca²⁺, leading to an increase in the level of another second messenger, cyclic guanosine monophosphate (cGMP). Cyclic GMP then binds to a protein kinase, stimulating a protein kinase cascade that leads to muscle relaxation. So far, the pathway seemed to conform to what was generally understood about signal transduction in general.

While this signal transduction pathway seemed to work in intact animals, it did not work on isolated strips of artery tissue. However, when Furchtgott switched to tubular sections of artery, signal transduction did occur. What accounted for the different results between tissue strips and tubular sections? Furchtgott realized that the endothelium, the delicate inner layer of cells lining the blood vessels, was lost during preparation of

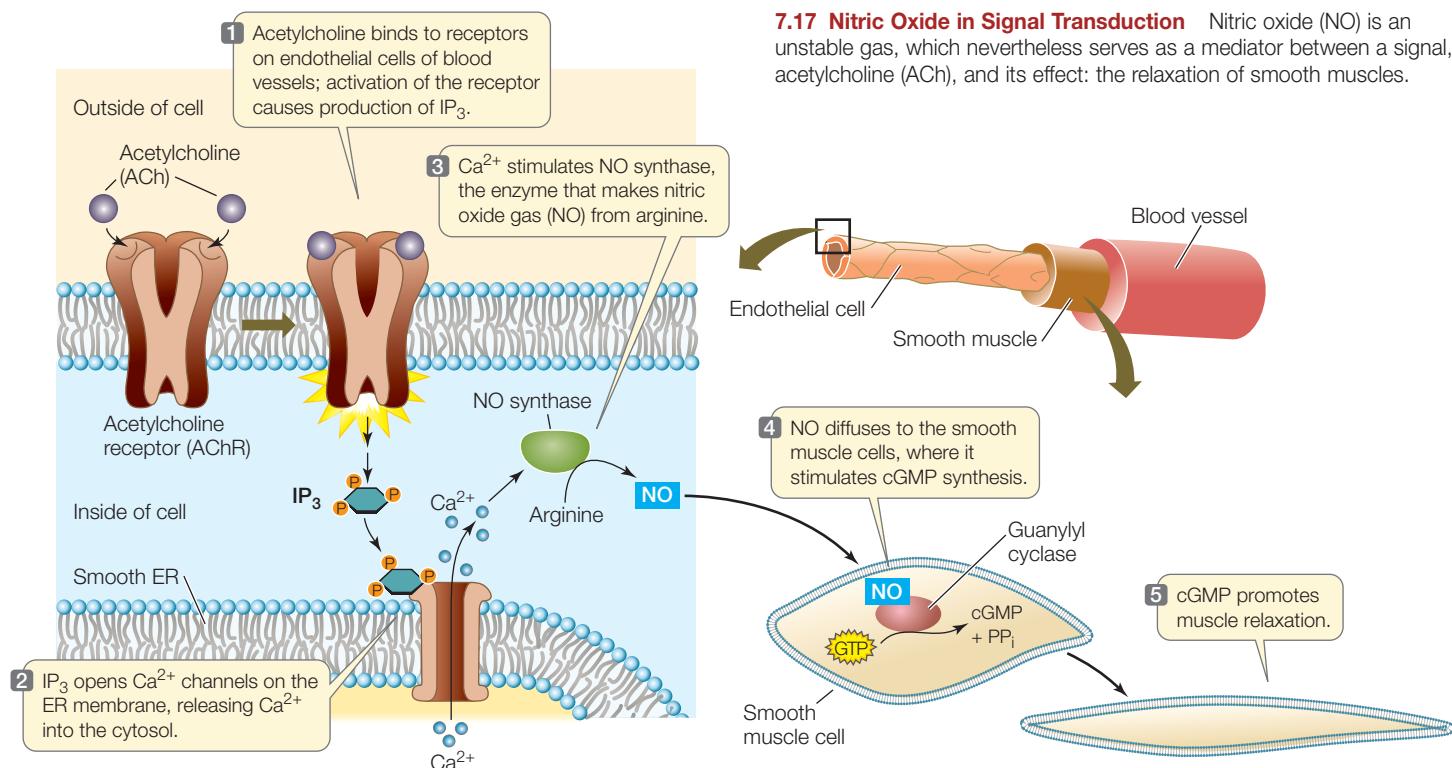


7.16 Calcium Ions as Second Messengers The concentration of Ca²⁺ can be measured using a dye that fluoresces when it binds the ion. Here, fertilization in a starfish egg causes a rush of Ca²⁺ from the environment into the cytoplasm. Areas of high calcium ion concentration are indicated by the red color and the events are photographed at 5-second intervals. Calcium signaling occurs in virtually all animal groups and triggers cell division in fertilized eggs, initiating the development of new individuals.

the tissue strips. He hypothesized that the endothelium was producing some chemical that diffused into the smooth muscle cells and was needed for their response to acetylcholine. However, the substance was not easy to isolate. It seemed to break down quickly, with a half-life (the time in which half of it disappeared) of 5 seconds in living tissue.

Furchtgott’s elusive substance turned out to be a gas, **nitric oxide (NO)**, which formerly had been recognized only as a toxic air pollutant! In the body, NO is made from the amino acid arginine by the enzyme NO synthase. When the acetylcholine receptor on the surface of an endothelial cell is activated, IP₃ is released, causing a calcium channel on the ER membrane to open and a subsequent increase in cytosolic Ca²⁺. The Ca²⁺ then activates NO synthase to produce NO. NO is chemically very unstable, readily reacting with oxygen gas as well as other small molecules. Although NO diffuses readily, it does not get far. Conveniently, the endothelial cells are close to the smooth muscle cells, where NO acts as a paracrine signal. In smooth muscle, NO activates an enzyme called guanylyl cyclase, catalyzing the formation of cGMP, which in turn relaxes the muscle cells (Figure 7.17).

The discovery of NO as a participant in signal transduction explained the action of nitroglycerin, a drug that has been used for over a century to treat angina, the chest pain caused by insufficient blood flow to the heart. Nitroglycerin releases NO, which results in relaxation of the blood vessels and increased blood flow. The drug sildenafil (Viagra) was developed to treat angina via the NO signal transduction pathway, but was only modestly useful for that purpose. However, men taking it reported more pronounced penile erections. During sexual stimulation, NO acts as a signal causing an increase in cGMP and a subsequent relaxation of the smooth muscles surrounding the arteries in the corpus cavernosum of the penis. As a result of this signal, the penis



fills with blood, producing an erection. Sildenafil acts by inhibiting an enzyme (a phosphodiesterase) that breaks down cGMP—resulting in more cGMP and better erections.

Signal transduction is highly regulated

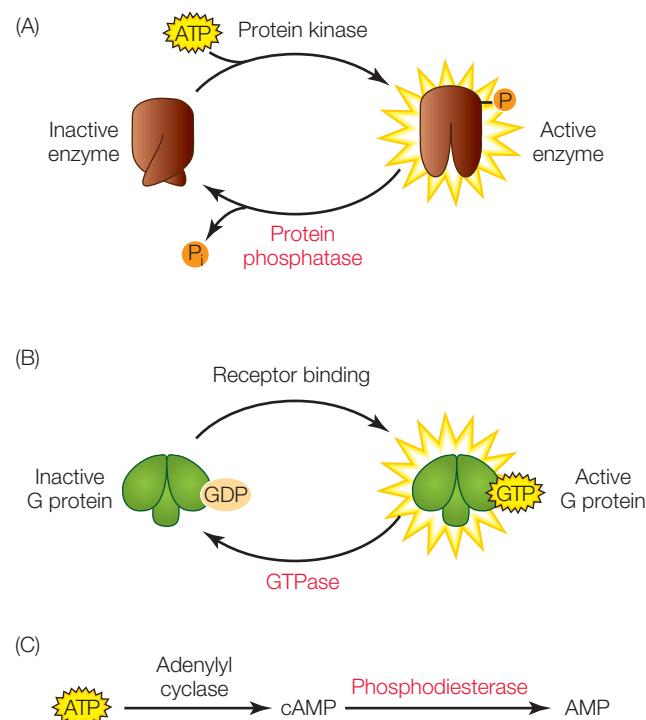
There are several ways in which cells can regulate the activity of a transducer. The concentration of NO, which breaks down quickly, can be regulated only by how much of it is made. On the other hand, membrane pumps and ion channels regulate the concentration of Ca²⁺, as we have seen. To regulate protein kinase cascades, G proteins, and cAMP, there are enzymes that convert the activated transducer back to its inactive precursor (**Figure 7.18**).

The balance between the activities of enzymes that activate transducers (for example, protein kinase) and enzymes that inactivate them (for example, protein phosphatase) is what determines the ultimate cellular response to a signal. Cells can alter this balance in several ways:

- **Synthesis or breakdown of the enzymes involved.** For example, synthesis of adenylyl cyclase and breakdown of phosphodiesterase (which breaks down cAMP) would tilt the balance in favor of more cAMP in the cell.
- **Activation or inhibition of the enzymes by other molecules.** Examples include the activation of a G protein-linked receptor by ligand binding, and inhibition of phosphodiesterase (which also breaks down cAMP) by sildenafil.

Because cell signaling is so important in diseases such as cancer, a search is under way for new drugs that can modulate the activities of enzymes that participate signal transduction pathways.

7.17 Nitric Oxide in Signal Transduction Nitric oxide (NO) is an unstable gas, which nevertheless serves as a mediator between a signal, acetylcholine (ACh), and its effect: the relaxation of smooth muscles.

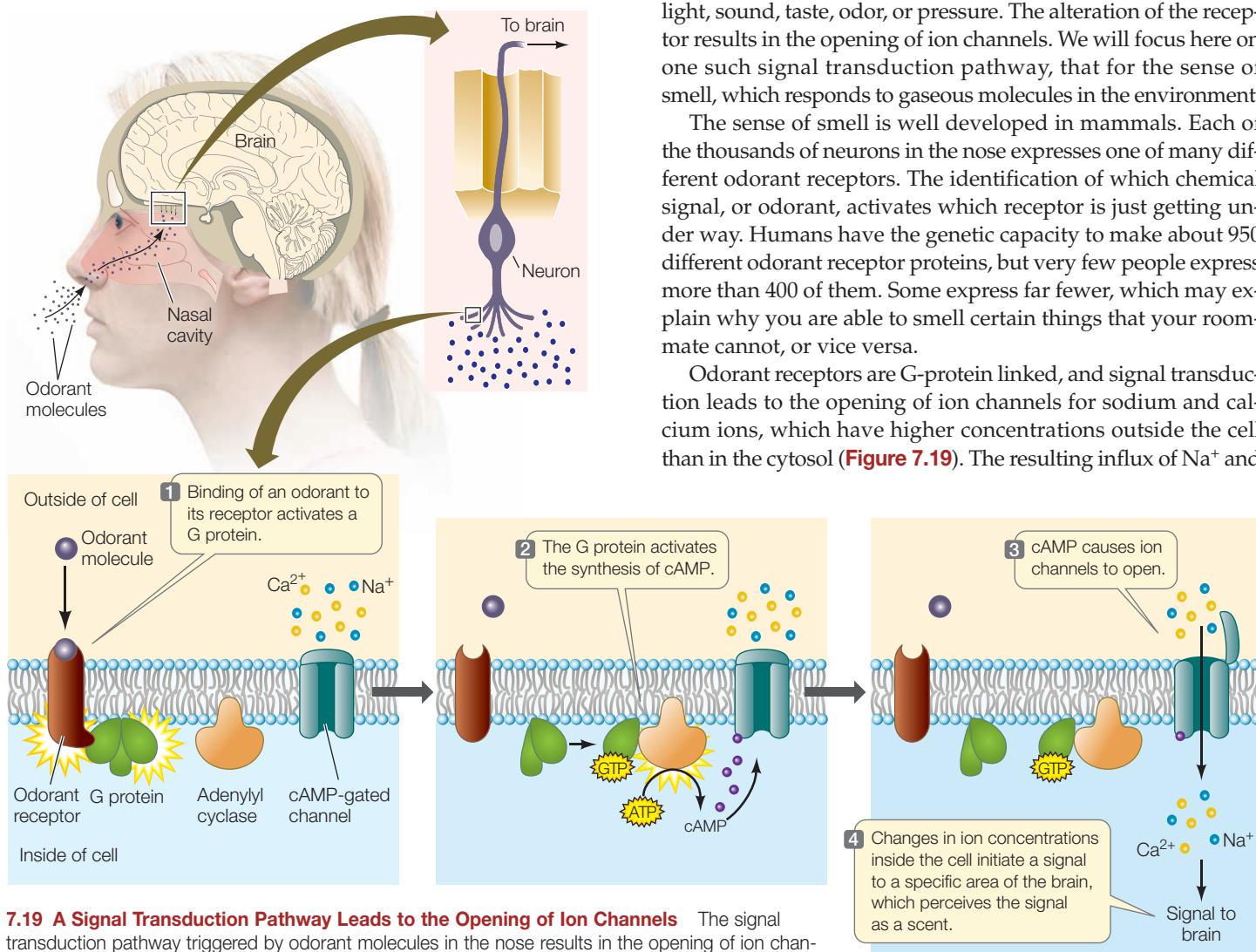


7.18 Regulation of Signal Transduction Some signals lead to the production of active transducers such as (A) protein kinases, (B) G proteins, and (C) cAMP. Other enzymes (red type) inactivate or remove these transducers.

7.3 RECAP

Signal transduction is the series of steps between the binding of a signal to a receptor and the ultimate cellular response. A receptor can activate a signal transduction pathway, such as a protein kinase cascade, directly. In many cases, a second messenger serves to amplify the signal and activate the signaling pathway indirectly. Protein kinase cascades amplify, distribute, and regulate signaling.

- How does a protein kinase cascade amplify a signal's message inside the cell? See pp. 136–137 and Figure 7.12
- What is the role of cAMP as a second messenger? See p. 138
- How are signal transduction cascades regulated? See p. 141 and Figure 7.18



7.19 A Signal Transduction Pathway Leads to the Opening of Ion Channels The signal transduction pathway triggered by odorant molecules in the nose results in the opening of ion channels. The resulting influx of Na^+ and Ca^{2+} into the neuron cells of the nose stimulates the transmission of a scent message to a specific region of the brain.

We have seen how the binding of a signal to its receptor initiates the response of a cell to the signal, and how signal transduction pathways amplify the signal and distribute its effects to numerous targets in the cell. In the next section we will consider the third step in the signal transduction process, the actual effects of the signal on cell function.

7.4 How Do Cells Change in Response to Signals?

The effects of a signal on cell function take three primary forms: the opening of ion channels, changes in the activities of enzymes, or differential gene expression. These events set the cell on a path for further and sometimes dramatic changes in form and function.

Ion channels open in response to signals

The opening of ion channels is a key step in the response of the nervous system to signals. In the sense organs, specialized cells have receptors that respond to external stimuli such as light, sound, taste, odor, or pressure. The alteration of the receptor results in the opening of ion channels. We will focus here on one such signal transduction pathway, that for the sense of smell, which responds to gaseous molecules in the environment.

The sense of smell is well developed in mammals. Each of the thousands of neurons in the nose expresses one of many different odorant receptors. The identification of which chemical signal, or odorant, activates which receptor is just getting under way. Humans have the genetic capacity to make about 950 different odorant receptor proteins, but very few people express more than 400 of them. Some express far fewer, which may explain why you are able to smell certain things that your roommate cannot, or vice versa.

Odorant receptors are G-protein linked, and signal transduction leads to the opening of ion channels for sodium and calcium ions, which have higher concentrations outside the cell than in the cytosol (Figure 7.19). The resulting influx of Na^+ and

Ca^{2+} causes the neuron to become stimulated so that it sends a signal to the brain that a particular odor is present.

Enzyme activities change in response to signals

Proteins will change their shapes if they are modified either covalently or noncovalently. We have seen examples of both types of modification in our description of signal transduction. A protein kinase adds a phosphate group to a target protein, and this covalent change alters the protein's conformation and activates or inhibits a function. Cyclic AMP binds noncovalently to a target protein, and this changes the protein's shape, activating or inhibiting its function. In the case of activation, a previously inaccessible active site is exposed, and the target protein goes on to perform a new cellular role.

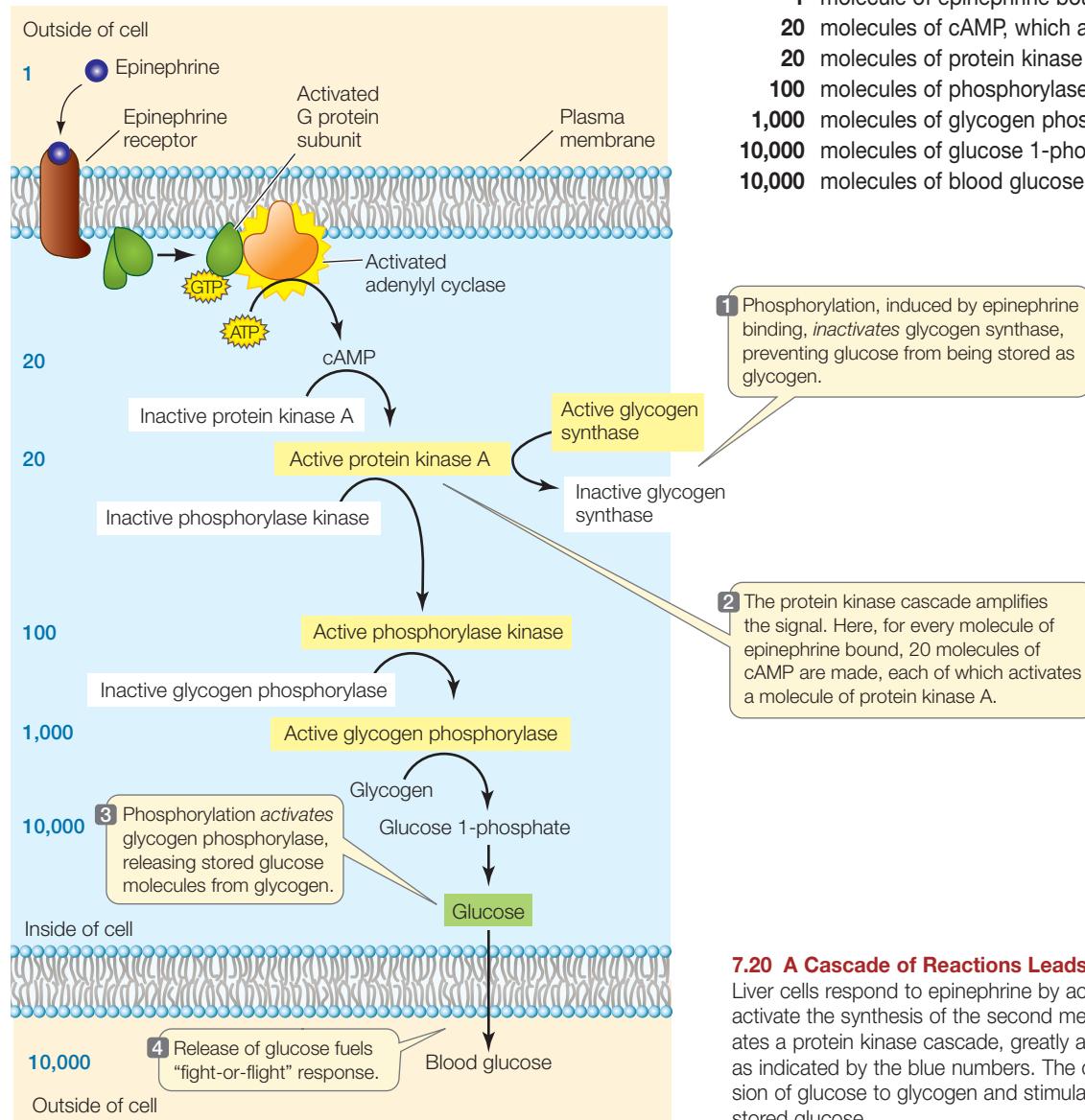
The G protein-mediated protein kinase cascade that is stimulated by epinephrine in liver cells results in the activation by cAMP of a key signaling molecule, protein kinase A. In turn,

protein kinase A phosphorylates two other enzymes, with opposite effects:

- **Inhibition.** Glycogen synthase, which catalyzes the joining of glucose molecules to synthesize the energy-storing molecule glycogen, is inactivated when a phosphate group is added to it by protein kinase A. Thus the epinephrine signal *prevents glucose from being stored in glycogen* (**Figure 7.20, step 1**).
- **Activation.** Phosphorylase kinase is activated when a phosphate group is added to it. It is part of a protein kinase cascade that ultimately leads to the activation of glycogen phosphorylase, another key enzyme in glucose metabolism. This enzyme results in the *liberation of glucose molecules from glycogen* (**Figure 7.20, steps 2 and 3**).

The amplification of the signal in this pathway is impressive; as detailed in Figure 7.20, each molecule of epinephrine that arrives at the plasma membrane ultimately results in 10,000 molecules of blood glucose:

- | |
|--|
| 1 molecule of epinephrine bound to the membrane activates
20 molecules of cAMP, which activate
20 molecules of protein kinase A, which activate
100 molecules of phosphorylase kinase, which activate
1,000 molecules of glycogen phosphorylase, which produce
10,000 molecules of glucose 1-phosphate, which produce
10,000 molecules of blood glucose |
|--|



7.20 A Cascade of Reactions Leads to Altered Enzyme Activity

Liver cells respond to epinephrine by activating G proteins, which in turn activate the synthesis of the second messenger cAMP. Cyclic AMP initiates a protein kinase cascade, greatly amplifying the epinephrine signal, as indicated by the blue numbers. The cascade both inhibits the conversion of glucose to glycogen and stimulates the release of previously stored glucose.

Signals can initiate DNA transcription

As we introduce in Section 4.1, the genetic material, DNA, is expressed by transcription as RNA, which is then translated into a protein whose amino acid sequence is specified by the original DNA sequence. Proteins are important in all cellular functions, so a key way to regulate specific functions in a cell is to regulate which proteins are made, and therefore, which DNA sequences are transcribed.

Signal transduction plays an important role in determining which DNA sequences are transcribed. Common targets of signal transduction are proteins called transcription factors, which bind to specific DNA sequences in the cell nucleus and activate or inactivate transcription of the adjacent DNA regions. For example, the Ras signaling pathway ends in the nucleus (see Figure 7.12). The final protein kinase in the Ras signaling cascade, MAPk, enters the nucleus and phosphorylates a protein which stimulates the expression of a number of genes involved in cell proliferation.

In this chapter we have concentrated on signaling pathways that occur in animal cells. However, as you will see in Part Eight of this book, plants also have signal transduction pathways, with equally important roles.

7.4 RECAP

Cells respond to signal transduction by activating enzymes, opening membrane channels, or initiating gene transcription.

- What role does cAMP play in the sense of smell? See pp. 142–143 and Figure 7.19
- How does amplification of a signal occur and why is it important in a cell’s response to changes in its environment? See p. 143 and Figure 7.20

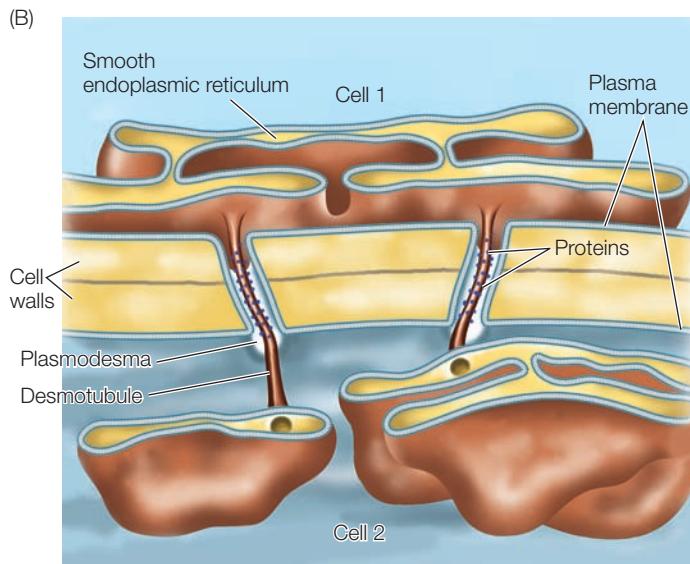
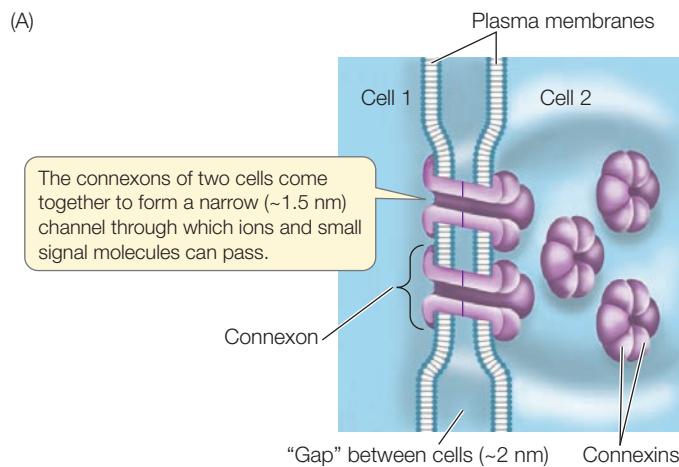
We have described how signals from a cell’s environment can influence the cell. But the environment of a cell in a multicellular organism is more than the extracellular medium—it includes neighboring cells as well. In the next section we’ll look at specialized junctions between cells that allow them to signal one another directly.

7.5 How Do Cells Communicate Directly?

Most cells are in contact with their neighbors. Section 6.2 describes various ways in which cells adhere to one another, such as via recognition proteins that protrude from the cell surface, or via tight junctions and desmosomes. But as we know from our own experience with our neighbors (and roommates), just being in proximity does not necessarily mean that there is functional communication. Neither tight junctions nor desmosomes are specialized for intercellular communication. However, many multicellular organisms have specialized cell junctions that allow their cells to communicate directly. In animals, these structures are gap junctions; in plants, they are plasmodesmata.

Animal cells communicate by gap junctions

Gap junctions are channels between adjacent cells that occur in many animals, occupying up to 25 percent of the area of the plasma membrane (Figure 7.21A). Gap junctions traverse the narrow space between the plasma membranes of two cells (the “gap”) by means of channel structures called **connexons**. The walls of a connexon are composed of six subunits of the integral membrane protein connexin. In adjacent cells, two connexons come together to form a gap junction that links the cytoplasms of the two cells. There may be hundreds of these channels between a cell and its neighbors. The channel pores are about 1.5 nm in diameter—far too narrow for the passage of large molecules such as proteins. But they are wide enough to allow small mol-



7.21 Communicating Junctions (A) An animal cell may contain hundreds of gap junctions connecting it to neighboring cells. The pores of gap junctions allow small molecules to pass from cell to cell, assuring similar concentrations of important signaling molecules in adjacent cells so that the cells can carry out the same functions. (B) Plasmodesmata connect plant cells. The desmotubule, derived from the smooth endoplasmic reticulum, fills up most of the space inside a plasmodesma, leaving a tiny gap through which small metabolites and ions can pass.

ecules to pass between the cells. Experiments in which labeled signal molecules or ions are injected into one cell show that they can readily pass into adjacent cells if the cells are connected by gap junctions. Why is it necessary to have these linkages between the cytoplasms of adjacent cells?

Gap junctions permit *metabolic cooperation* between the linked cells. Such cooperation ensures the sharing between cells of important small molecules such as ATP, metabolic intermediates, amino acids, and coenzymes (see Section 8.4). In some tissues, metabolic cooperation is needed so that signals and metabolic products can be passed from cells at the edges of tissues to cells in the interior and vice versa. It is not clear how important this function is in many tissues, but it is known to be vital in some. For example, in the lens of the mammalian eye only the cells at the periphery are close enough to the blood supply to allow diffusion of nutrients and wastes. But because lens cells are connected by large numbers of gap junctions, material can diffuse between them rapidly and efficiently.

As mentioned above, there is evidence that signal molecules such as hormones and second messengers such as cAMP can move through gap junctions. If this is true, then only a few cells would need receptors for a signal in order for the signal to be transduced throughout the tissue. In this way, a tissue can have a coordinated response to the signal.

Plant cells communicate by plasmodesmata

Instead of gap junctions, plants have **plasmodesmata** (singular *plasmodesma*), which are membrane-lined tunnels that traverse the thick cell walls separating plant cells from one another. A typical plant cell has several thousand plasmodesmata.

Plasmodesmata differ from gap junctions in one fundamental way: unlike gap junctions, in which the wall of the channel is made of integral proteins from the adjacent plasma membranes,

plasmodesmata are lined by the fused plasma membranes themselves. Plant biologists are so familiar with the notion of a tissue as cells interconnected in this way that they refer to these continuous cytoplasms as a *symplast* (see Figure 35.6).

The diameter of a plasmodesma is about 6 nm, far larger than a gap junction channel. But the actual space available for diffusion is about the same—1.5 nm. Examination of the interior of the plasmodesma by transmission electron microscopy reveals that a tubule called the **desmotubule**, apparently derived from the endoplasmic reticulum, fills up most of the opening of the plasmodesma (Figure 7.21B). Typically, only small metabolites and ions can move between plant cells. This fact is important in plant physiology because the bulk transport system in plants, the vascular system, lacks the tiny circulatory vessels (capillaries) that many animals have for bringing gases and nutrients to every cell. Diffusion from cell to cell across plasma membranes is probably inadequate to account for the movement of a plant hormone from the site of production to the site of action. Instead, plants rely on more rapid diffusion through plasmodesmata to ensure that all cells of a tissue respond to a signal at the same time. There are cases in which larger molecules or particles can pass between cells via plasmodesmata. For example, some viruses can move through plasmodesmata by using “movement proteins” to assist their passage.

7.5 RECAP

Cells can communicate with their neighbors through specialized cell junctions. In animals, these structures are gap junctions; in plants, they are plasmodesmata.

- What are the roles that gap junctions and plasmodesmata play in cell signaling?

CHAPTER SUMMARY

7.1 What Are Signals, and How Do Cells Respond to Them?

- Cells receive many signals from the physical environment and from other cells. Chemical signals are often at very low concentrations. **Autocrine** signals affect the cells that make them; **paracrine** signals diffuse to and affect nearby cells. **Review Figure 7.1, WEB ACTIVITY 7.1**
- A **signal transduction pathway** involves the interaction of a signal molecule with a **receptor**; the transduction and amplification of the signal via a series of steps within the cell; and effects on the function of the cell. **Review Figure 7.2**

7.2 How Do Signal Receptors Initiate a Cellular Response?

- Cells respond to signals only if they have specific receptor proteins that can bind those signals. Depending on the nature of its signal or **ligand**, a receptor may be located in the plasma membrane or in the cytoplasm of the target cell. **Review Figure 7.5**
- Receptors located in the plasma membrane include **ion channels**, **protein kinases**, and **G protein-linked receptors**.

- Ion channel receptors are “gated”: the gate “opens” when the three-dimensional structure of the channel protein is altered by ligand binding. **Review Figure 7.6**

- A **G protein** has three important binding sites, which bind a G protein-linked receptor, GDP or GTP, and an **effector protein**. A G protein can either activate or inhibit an effector protein. **Review Figure 7.8, ANIMATED TUTORIAL 7.1**

- Lipid-soluble signals, such as steroid hormones, can diffuse through the plasma membrane and meet their receptors in the cytoplasm; the ligand–receptor complex may then enter the nucleus to affect gene expression. **Review Figure 7.9**

7.3 How Is the Response to a Signal Transduced through the Cell?

- **Direct signal transduction** is a function of the receptor itself and occurs at the plasma membrane. **Indirect transduction** involves a soluble **second messenger**. **Review Figure 7.10**
- A **protein kinase cascade** amplifies the response to receptor binding. **Review Figure 7.12, ANIMATED TUTORIAL 7.2**

- Second messengers include **cyclic AMP (cAMP)**, **inositol trisphosphate (IP₃)**, **diacylglycerol (DAG)**, and **calcium ions**. IP₃ and DAG are derived from the phospholipid **phosphatidyl inositol-bisphosphate (PIP2)**.
- The gas **nitric oxide (NO)** is involved in signal transduction in human smooth muscle cells. **Review Figure 7.17**
- Signal transduction can be regulated in several ways. The balance between activating and inactivating the molecules involved determines the ultimate cellular response to a signal. **Review Figure 7.18**

7.4 How Do Cells Change in Response to Signals?

- The cellular responses to signals may be the opening of ion channels, the alteration of enzyme activities, or changes in gene expression. **Review Figure 7.19**
- Protein kinases covalently add phosphate groups to target proteins; cAMP binds target proteins noncovalently. Both kinds of

binding change the target protein's conformation to expose or hide its active site.

- Activated enzymes may activate other enzymes in a signal transduction pathway, leading to impressive amplification of a signal. **Review Figure 7.20**

7.5 How Do Cells Communicate Directly?

- Many adjacent animal cells can communicate with one another directly through small pores in their plasma membranes called **gap junctions**. Protein structures called **connexons** form thin channels between two adjacent cells through which small signal molecules and ions can pass. **Review Figure 7.21A**
- Plant cells are connected by somewhat larger pores called **plasmodesmata**, which traverse both plasma membranes and cell walls. The **desmotubule** narrows the opening of the plasmodesma. **Review Figure 7.21B**

SEE WEB ACTIVITY 7.2 for a concept review of this chapter.

SELF-QUIZ

- What is the correct order for the following events in the interaction of a cell with a signal? (1) Alteration of cell function; (2) signal binds to receptor; (3) signal released from source; (4) signal transduction.
 - 1234
 - 2314
 - 3214
 - 3241
 - 3421
- Why do some signals ("first messengers") trigger "second messengers" to activate target cells?
 - The first messenger requires activation by ATP.
 - The first messenger is not water soluble.
 - The first messenger binds to many types of cells.
 - The first messenger cannot cross the plasma membrane.
 - There are no receptors for the first messenger.
- Steroid hormones such as estrogen act on target cells by
 - initiating second messenger activity.
 - binding to membrane proteins.
 - initiating gene expression.
 - activating enzymes.
 - binding to membrane lipids.
- The major difference between a cell that responds to a signal and one that does not is the presence of a
 - DNA sequence that binds to the signal.
 - nearby blood vessel.
 - receptor.
 - second messenger.
 - transduction pathway.
- Which of the following is *not* a consequence of a signal binding to a receptor?
 - Activation of receptor enzyme activity
 - Diffusion of the receptor in the plasma membrane
 - Change in conformation of the receptor protein
 - Breakdown of the receptor to amino acids
 - Release of the signal from the receptor
- A nonpolar molecule such as a steroid hormone usually binds to a
 - cytoplasmic receptor.
 - protein kinase.
 - ion channel.
 - phospholipid.
 - second messenger.
- Which of the following is *not* a common type of receptor?
 - Ion channel
 - Protein kinase
 - G protein-linked receptor
 - Cytoplasmic receptor
 - Adenyllyl cyclase
- Which of the following is *not* true of a protein kinase cascade?
 - The signal is amplified.
 - A second messenger is formed.
 - Target proteins are phosphorylated.
 - The cascade ends up at the mitochondrion.
 - The cascade begins at the plasma membrane.
- Which of the following is *not* a second messenger?
 - Calcium ion
 - Inositol trisphosphate
 - ATP
 - Cyclic AMP
 - Diacylglycerol
- Plasmodesmata and gap junctions
 - allow small molecules and ions to pass rapidly between cells.
 - are both membrane-lined channels.
 - are channels about 1 mm in diameter.
 - are present only once per cell.
 - are involved in cell recognition.

FOR DISCUSSION

1. Like the Ras protein itself, the various components of the Ras signaling pathway were discovered when cancer cells showed changes (mutations) in the genes encoding one or another of the components. What might be the biochemical consequences of mutations in the genes coding for (a) Raf and (b) MAP kinase that resulted in rapid cell division?
2. Cyclic AMP is a second messenger in many different responses. How can the same messenger act in different ways in different cells?
3. Compare direct communication via plasmodesmata or gap junctions with receptor-mediated communication between cells. What are the advantages of one method over the other?
4. The tiny invertebrate Hydra has an apical region with tentacles and a long, slender body. Hydra can reproduce asexually when cells on the body wall differentiate and form a bud, which then breaks off as a new organism. Buds form only at certain distances from the apex, leading to the idea that the apex releases a signal molecule that diffuses down the body and, at high concentrations (i.e., near the apex), inhibits bud formation. Hydra lacks a circulatory system, so this inhibitor must diffuse from cell to cell. If you had an antibody that binds to connexons and plugs up the gap junctions, how would you test the hypothesis that Hydra's inhibitory factor passes through these junctions?

ADDITIONAL INVESTIGATION

Endosymbiotic bacteria in the marine invertebrate *Begula neritina* synthesize bryostatins, a name derived from the invertebrate's animal group Ectoprocta, once known as bryozoans ("moss animals"), and *stat* (stop). When used as drugs, bryostatins curtail

cell division in many cell types, including several cancers. It has been proposed that bryostatins inhibit protein kinase C (see Figure 7.15). How would you investigate this hypothesis, and how would you relate this inhibition to cell division?

WORKING WITH DATA (GO TO yourBioPortal.com)

The Discovery of a Second Messenger In this hands-on exercise, you will examine the experiments that Sutherland and his colleagues performed (Figure 7.13) using liver tissue to demonstrate that there can be a second, soluble chemical

messenger between a hormone binding to a receptor and its eventual effects in the cell. By analyzing their data, you will see how controls were important in their reasoning.

Energy, Enzymes, and Metabolism

Lactase deficiency

United Nations officials first noticed the problem during the 1950s when massive food relief efforts were made to alleviate famines in Asia and Africa. The conventional wisdom was that donated food should provide a balanced diet, and that an important component of the diet (one of the “four major food groups”) was dairy products. Reports started coming in of people developing bloating, nausea, and diarrhea after consuming donated dairy products.

At first, this problem was attributed to contamination by bacteria during shipping, or to errors in the preparation of powdered milk products by the recipients. It never occurred to the donors that the scientific principles of nutrition that they had so carefully developed did not apply to people everywhere. But it soon became apparent that

the donors in Europe and other wealthy countries, who were usually of European descent, were atypical of humanity in their ability to hydrolyze the disaccharide lactose, the major milk sugar, to its constituent monosaccharides, glucose and galactose. Their small intestines make a protein called lactase (β -galactosidase) that acts to speed up the hydrolysis reaction millions-fold. Such catalytic proteins are called enzymes, and their names often end with the suffix “ase.” Most people around the world are born with the ability to make the enzyme lactase, but soon after infancy they lose it. People of European descent are unusual in that they do not lose their lactase production after infancy.

When many non-European adults consume lactose it does not get hydrolyzed in their small intestine, because they do not produce lactase. Disaccharides such as lactose are not absorbed into the blood stream by cells lining the small intestine. So the lactose remains intact and travels onward to the colon (large intestine). Among the billions of bacteria in the colon, there are species that make lactase. But as a side product, these bacteria produce the gases that cause all the discomfort. The condition of discomfort after eating lactose is called lactose intolerance.

Why does lactase production go down after infancy in most humans? The explanation lies with diet: an infant first consumes mother’s milk, which contains abundant lactose. This stimulates the intestinal cells to make lactase. But many humans—and other mammals—consume little or no milk after weaning, and the ability to make lactase in the small intestine is not needed. So most mammals have evolved to produce lactase only during infancy. Lactose intolerance is not a problem in many human societies because the people simply don’t consume dairy products—



A Precursor to Trouble Many adults do not produce the enzyme lactase in their small intestines. When they consume dairy products, these people have ill effects.



Maasai Herders The Maasai are unusual among Africans in that they consume milk throughout their lives. They can do this because they produce lactase after weaning.

unless they are given them by well-meaning donors! They get their carbohydrates from other sources.

Then why are many people of European descent still able to make lactase as adults? It turns out that they carry a mutation (a change in their DNA sequence) that eliminates the shutdown in lactase production after weaning. This mutation became predominant in European (and some east African) populations after those people began to keep grazing animals and to use their milk.

Lactase activity is an example of an enzyme-catalyzed biochemical transformation. The hydrolysis of lactose is the beginning of its transformation to simpler molecules—ultimately CO_2 —and this transformation releases energy.

IN THIS CHAPTER we begin our study of biochemical transformations, focusing on the role of energy. We first describe the physical principles that underlie energy transformations and how these principles apply to biology. Then, we go on to show how the energy carrier ATP plays an important role in the cell. Finally, we follow up on the lactase story by describing the nature, activities, and regulation of enzymes, which speed up biochemical transformations and are essential for life.

CHAPTER OUTLINE

- 8.1 What Physical Principles Underlie Biological Energy Transformations?
- 8.2 What Is the Role of ATP in Biochemical Energetics?
- 8.3 What Are Enzymes?
- 8.4 How Do Enzymes Work?
- 8.5 How Are Enzyme Activities Regulated?

8.1 What Physical Principles Underlie Biological Energy Transformations?

Metabolic reactions and catalysts are essential to the biochemical transformation of energy by living things. Whether it is a plant using light energy to produce carbohydrates or a cat transforming food energy so it can leap to a countertop (where it hopes to find food so it can obtain more energy), the transformation of energy is a hallmark of life.

Physicists define energy as the capacity to do work. Work occurs when a force operates on an object over a distance. In biochemistry, it is more useful to consider energy as *the capacity for change*. In biochemical reactions these energy changes are usually associated with changes in the chemical composition and properties of molecules. No cell creates energy; all living things must obtain energy from the environment. Indeed one of the fundamental laws of physics is that energy can neither be created nor destroyed. However, energy can be transformed from one form into another, and living cells carry out many such transformations. For example, green plant cells convert light energy into chemical energy; the jumping cat transforms chemical energy into movement. Energy transformations are linked to the chemical transformations that occur in cells—the breaking and creating of chemical bonds, the movement of substances across membranes, cell reproduction, and so forth.

There are two basic types of energy and of metabolism

Energy comes in many forms: chemical, electrical, heat, light, and mechanical. But all forms of energy can be considered as one of two basic types:

- *Potential energy* is the energy of state or position—that is, stored energy. It can be stored in many forms: in chemical bonds, as a concentration gradient, or even as an electric charge imbalance (as in the membrane potential; see Section 6.3). Think of a crouching cat, holding still as it prepares to pounce.
- *Kinetic energy* is the energy of movement—that is, the type of energy that does work, that makes things change. Think of the cat leaping as some of the potential energy stored in its muscles is converted into the kinetic energy of muscle contractions.

Potential energy can be converted into kinetic energy and vice versa, and the form that the energy takes can also be converted. The potential energy in the cat's muscles is in covalent bonds (chemical energy), while the kinetic energy of the pouncing cat is mechanical (**Figure 8.1**). You can think of many other such



8.1 Energy Conversions and Work A leaping cat illustrates both the conversion between potential and kinetic energy and the conversion of energy from one form (chemical) to another (mechanical).

conversions: while reading this book, for example, light energy is converted into chemical energy in your eyes, and then is converted into electric energy in the nerve cells that carry messages to your brain. When you decide to turn a page, the electrical and chemical energy of nerve and muscle are converted into kinetic energy.

In any living organism, chemical reactions are occurring continuously. **Metabolism** is defined as the totality of these reactions. While particular cells carry out many reactions at any given instant, scientists usually focus on a few reactions at a time. Two broad categories of metabolic reactions occur in all cells of all organisms:

- **Anabolic reactions** (anabolism) link simple molecules to form more complex molecules (for example, the synthesis of a protein from amino acids). Anabolic reactions require an input of energy and capture it in the chemical bonds that are formed.
- **Catabolic reactions** (catabolism) break down complex molecules into simpler ones and release the energy stored in chemical bonds. For example, when the polysaccharide starch is hydrolyzed to simpler molecules, energy is released.

Catabolic and anabolic reactions are often linked. The energy released in catabolic reactions is often used to drive anabolic re-

actions—that is, to do biological work. For example, the energy released by the breakdown of glucose (catabolism) is used to drive anabolic reactions such as the synthesis of nucleic acids and proteins.

Catabolic reactions also provide energy for movement: muscle contraction is driven by the catabolism (hydrolysis) of ATP (see Section 8.2). In this case, the potential energy released by catabolism is converted to kinetic energy.

The **laws of thermodynamics** (thermo, “energy”; dynamics, “change”) were derived from studies of the fundamental physical properties of energy, and the ways it interacts with matter. The laws apply to all matter and all energy transformations in the universe. Their application to living systems helps us to understand how organisms and cells harvest and transform energy to sustain life.

The first law of thermodynamics: Energy is neither created nor destroyed

The first law of thermodynamics states that in any conversion of energy, it is neither created nor destroyed. Another way of saying this is: in any conversion of energy, the total energy before and after the conversion is the same (**Figure 8.2A**). As you will see in the next two chapters, the potential energy present in the chemical bonds of carbohydrates and lipids can be converted to potential energy in the form of ATP. This can then be converted into kinetic energy to do mechanical work (such as in muscle contractions), or used to do biochemical work (such as protein synthesis).

The second law of thermodynamics: Disorder tends to increase

Although energy cannot be created or destroyed, the second law of thermodynamics states that when energy is converted from one form to another, some of that energy becomes unavailable for doing work (**Figure 8.2B**). In other words, no physical process or chemical reaction is 100 percent efficient; some of the released energy is lost to a form associated with disorder. Think of disorder as a kind of randomness due to the thermal motion of particles; this energy is of such a low value and so dispersed that it is unusable. *Entropy* is a measure of the disorder in a system.

It takes energy to impose order on a system. Unless energy is applied to a system, it will be randomly arranged or disordered. The second law applies to all energy transformations, but we will focus here on chemical reactions in living systems.

NOT ALL ENERGY CAN BE USED In any system, the total energy includes the usable energy that can do work and the unusable energy that is lost to disorder:

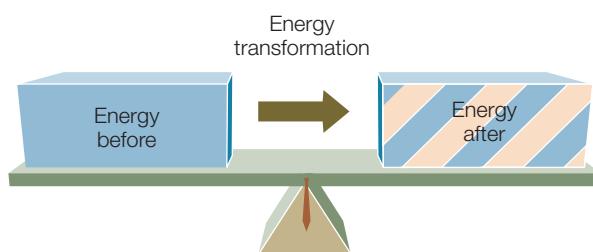
$$\text{total energy} = \text{usable energy} + \text{unusable energy}$$

In biological systems, the total energy is called **enthalpy (H)**. The usable energy that can do work is called **free energy (G)**. Free energy is what cells require for all the chemical reactions needed for growth, cell division, and maintenance. The unusable en-

(A)

The First Law of Thermodynamics

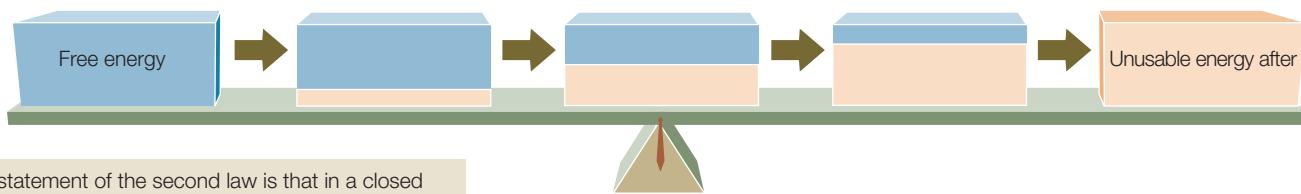
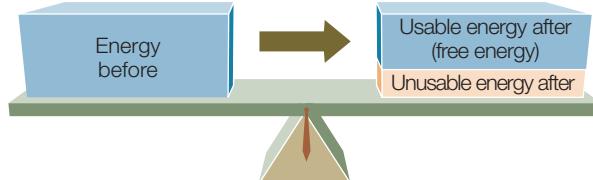
The total amount of energy before a transformation equals the total amount after a transformation. No new energy is created, and no energy is lost.



(B)

The Second Law of Thermodynamics

Although a transformation does not change the total amount of energy within a closed system (one that is not exchanging matter or energy with the surroundings), after any transformation the amount of energy available to do work is always less than the original amount of energy.



Another statement of the second law is that in a closed system, with repeated energy transformations, free energy decreases and unusable energy (disorder) increases—a phenomenon known as the increase in **entropy**.

ergy is represented by **entropy (S)** multiplied by the absolute temperature (T). Thus we can rewrite the word equation above more precisely as:

$$H = G + TS$$

Because we are interested in usable energy, we rearrange this expression:

$$G = H - TS$$

Although we cannot measure G , H , or S absolutely, we can determine the change in each at a constant temperature. Such energy changes are measured in calories (cal) or joules (J).* A change in energy is represented by the Greek letter delta (Δ). The change in free energy (ΔG) of any chemical reaction is equal to the difference in free energy between the products and the reactants:

$$\Delta G_{\text{reaction}} = G_{\text{products}} - G_{\text{reactants}}$$

Such a change can be either positive or negative; that is, the free energy of the products can be more or less than the free energy of the reactants. If the products have more free energy than the reactants, then there must have been some input of energy

*A calorie is the amount of heat energy needed to raise the temperature of 1 gram of pure water from 14.5°C to 15.5°C. In the SI system, energy is measured in joules. 1 J = 0.239 cal; conversely, 1 cal = 4.184 J. Thus, for example, 486 cal = 2,033 J, or 2.033 kJ. Although defined here in terms of heat, the calorie and the joule are measures of any form of energy—mechanical, electrical, or chemical. When you compare data on energy, always compare joules with joules and calories with calories.

8.2 The Laws of Thermodynamics

(A) The first law states that energy cannot be created or destroyed. (B) The second law states that after energy transformations, some energy becomes unavailable to do work.

into the reaction. (Remember that energy cannot be created, so some energy must have been added from an external source.) At a constant temperature, ΔG is defined in terms of the change in total energy (ΔH) and the change in entropy (ΔS):

$$\Delta G = \Delta H - T\Delta S$$

This equation tells us whether free energy is released or consumed by a chemical reaction:

- If ΔG is negative ($\Delta G < 0$), free energy is released.
- If ΔG is positive ($\Delta G > 0$), free energy is required (consumed).

If the necessary free energy is not available, the reaction does not occur. The sign and magnitude of ΔG depend on the two factors on the right of the equation:

- ΔH : In a chemical reaction, ΔH is the total amount of energy added to the system ($\Delta H > 0$) or released ($\Delta H < 0$).
- ΔS : Depending on the sign and magnitude of ΔS , the entire term, $T\Delta S$, may be negative or positive, large or small. In other words, in living systems at a constant temperature (no change in T), the magnitude and sign of ΔG can depend a lot on changes in entropy.

If a chemical reaction increases entropy, its products are more disordered or random than its reactants. If there are more products than reactants, as in the hydrolysis of a protein to its amino acids, the products have considerable freedom to move around. The disorder in a solution of amino acids will be large compared with that in the protein, in which peptide bonds and other forces prevent free movement. So in hydrolysis, the change in entropy (ΔS) will be positive. Conversely, if there are fewer products and they are more restrained in their movements than the reac-

tants (as for amino acids being joined in a protein), ΔS will be negative.

DISORDER TENDS TO INCREASE The second law of thermodynamics also predicts that, as a result of energy transformations, disorder tends to increase; some energy is always lost to random thermal motion (entropy). Chemical changes, physical changes, and biological processes all tend to increase entropy (see Figure 8.2B), and this tendency gives direction to these processes. It explains why some reactions proceed in one direction rather than another.

How does the second law apply to organisms? Consider the human body, with its highly organized tissues and organs composed of large, complex molecules. This level of complexity appears to be in conflict with the second law but is not for two reasons. First, the construction of complexity also generates disorder. Constructing 1 kg of a human body requires the metabolism of about 10 kg of highly ordered biological materials, which are converted into CO_2 , H_2O , and other simple molecules that move independently and randomly. So metabolism creates far more disorder (more energy is lost to entropy) than the amount of order (total energy; enthalpy) stored in 1 kg of flesh. Second, life requires a constant input of energy to maintain order. Without this energy, the complex structures of living systems would break down. Because energy is used to generate and maintain order, there is no conflict with the second law of thermodynamics.

Having seen that the laws of thermodynamics apply to living things, we will now turn to a consideration of how these laws apply to biochemical reactions.

Chemical reactions release or consume energy

Since anabolic reactions link simple molecules to form more complex molecules, they tend to increase complexity (order) in the cell. On the other hand, catabolic reactions break down complex molecules into simpler ones, so they tend to decrease complexity (generate disorder).

- *Catabolic* reactions may break down an ordered reactant into smaller, more randomly distributed products. Reactions that release free energy ($-\Delta G$) are called **exergonic** reactions (Figure 8.3A). For example:



- *Anabolic* reactions may make a single product (a highly ordered substance) out of many smaller reactants (less ordered). Reactions that require or consume free energy ($+\Delta G$) are called **endergonic** reactions (Figure 8.3B).

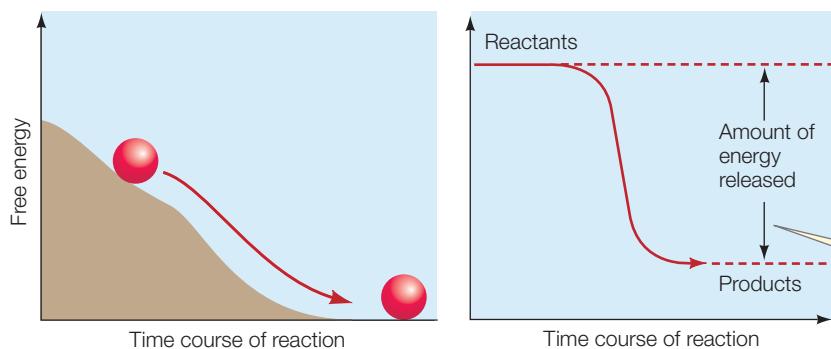
For example:



In principle, chemical reactions are reversible and can run both forward and backward. For example, if compound A can be converted into compound B ($A \rightarrow B$), then B, in principle, can be converted into A ($B \rightarrow A$), although *the concentrations of A and B determine which of these directions will be favored*. Think of the overall reaction as resulting from competition between forward and reverse reactions ($A \rightleftharpoons B$). Increasing the concentration of A speeds up the forward reaction, and increasing the concentration of B favors the reverse reaction.

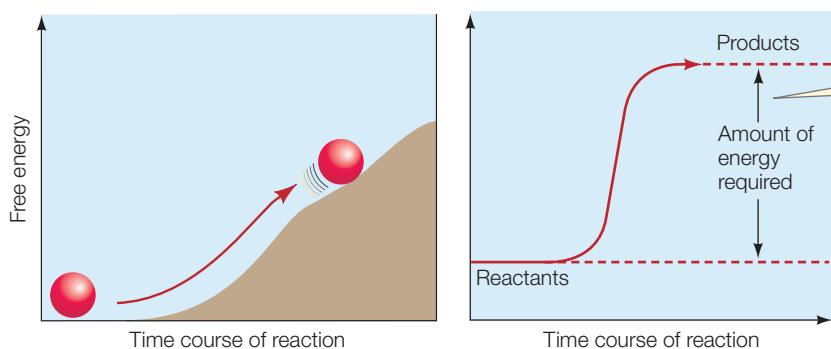
At some concentration of A and B, the forward and reverse reactions take place at the same rate. At this concentration, no further net change in the system is observable, although individual molecules are still forming and breaking apart. This balance between forward and reverse reactions is known as **chemical equilibrium**. Chemical equilibrium is a state of no net change, and a state in which $\Delta G = 0$.

(A) Exergonic reaction



In an exergonic reaction, energy is released as the reactants form lower-energy products. ΔG is negative.

(B) Endergonic reaction

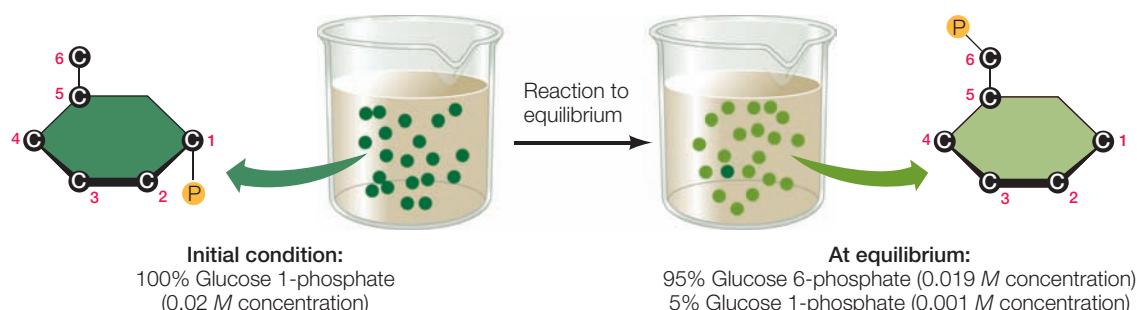


Energy must be added for an endergonic reaction, in which reactants are converted to products with a higher energy level. ΔG is positive.

8.3 Exergonic and Endergonic Reactions (A) In an exergonic reaction, the reactants behave like a ball rolling down a hill, and energy is released. (B) A ball will not roll uphill by itself. Driving an endergonic reaction, like moving a ball uphill, requires the addition of free energy.

8.4 Chemical Reactions Run to Equilibrium

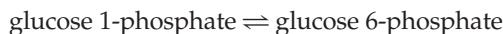
No matter what quantities of glucose 1-phosphate and glucose 6-phosphate are dissolved in water, when equilibrium is attained, there will always be 95 percent glucose 6-phosphate and 5 percent glucose 1-phosphate.



Chemical equilibrium and free energy are related

Every chemical reaction proceeds to a certain extent, but not necessarily to completion (all reactants converted into products). Each reaction has a specific equilibrium point, which is related to the free energy released by the reaction under specified conditions. To understand the principle of equilibrium, consider the following example.

Most cells contain glucose 1-phosphate, which is converted into glucose 6-phosphate.



Imagine that we start out with an aqueous solution of glucose 1-phosphate that has a concentration of 0.02 M. (M stands for molar concentration; see Section 2.4). The solution is maintained under constant environmental conditions (25°C and pH 7). As the reaction proceeds to equilibrium, the concentration of the product, glucose 6-phosphate, rises from 0 to 0.019 M, while the concentration of the reactant, glucose 1-phosphate, falls to 0.001 M. At this point, equilibrium is reached (**Figure 8.4**). At equilibrium, the reverse reaction, from glucose 6-phosphate to glucose 1-phosphate, progresses at the same rate as the forward reaction.

At equilibrium, then, this reaction has a product-to-reactant ratio of 19:1 (0.019/0.001), so the forward reaction has gone 95 percent of the way to completion ("to the right," as written above). This result is obtained every time the experiment is run under the same conditions.

The change in free energy (ΔG) for any reaction is related directly to its point of equilibrium. The further toward completion the point of equilibrium lies, the more free energy is released. In an exergonic reaction, such as the conversion of glucose 1-phosphate to glucose 6-phosphate, ΔG is a negative number (in this example, $\Delta G = -1.7 \text{ kcal/mol}$, or -7.1 kJ/mol).

A large, positive ΔG for a reaction means that it proceeds hardly at all to the right ($A \rightarrow B$). If the concentration of B is initially high relative to that of A, such a reaction runs "to the left" ($A \leftarrow B$), and nearly all B is converted into A. A ΔG value near zero is characteristic of a readily reversible reaction: reactants and products have almost the same free energies.

In Chapters 9 and 10 we examine the metabolic reactions that harvest energy from food and light. In turn, this energy is used to synthesize carbohydrates, lipids and proteins. All of the chemical reactions carried out by living organisms are governed by the principles of thermodynamics and equilibrium.

8.1 RECAP

Two laws of thermodynamics govern energy transformations in biological systems. A biochemical reaction can release or consume energy, and it may not run to completion, but instead end up at a point of equilibrium.

- What is the difference between potential energy and kinetic energy? Between anabolism and catabolism?
See pp. 149–150
- What are the laws of thermodynamics? How do they relate to biology? **See pp. 150–152 and Figure 8.2**
- What is the difference between endergonic and exergonic reactions and what is the importance of ΔG ?
See p. 152 and Figure 8.3

The principles of thermodynamics that we have been discussing apply to all energy transformations in the universe, so they are very powerful and useful. Next, we'll apply them to reactions in cells that involve the currency of biological energy, ATP.

8.2 What Is the Role of ATP in Biochemical Energetics?

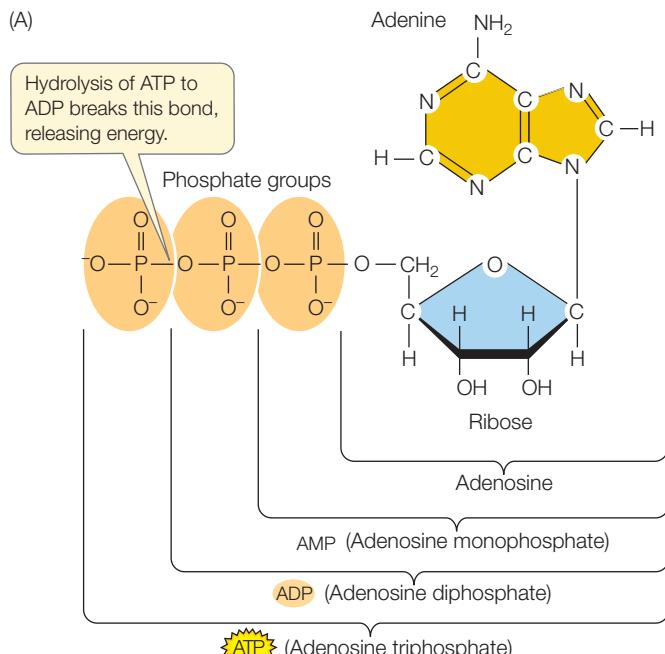
Cells rely on adenosine triphosphate (ATP) for the capture and transfer of the free energy they need to do chemical work. ATP operates as a kind of "energy currency." Just as it is more effective, efficient, and convenient for you to trade money for a lunch than to trade your actual labor, it is useful for cells to have a single currency for transferring energy between different reactions and cell processes. So some of the free energy that is released by exergonic reactions is captured in the formation of ATP from adenosine diphosphate (ADP) and inorganic phosphate (HPO_4^{2-} , which is commonly abbreviated to P_i). The ATP can then be hydrolyzed at other sites in the cell to release free energy to drive endergonic reactions. (In some reactions, guanosine triphosphate [GTP] is used as the energy transfer molecule instead of ATP, but we will focus on ATP here.)

ATP has another important role in the cell beyond its use as an energy currency: it is a nucleotide that can be converted into a building block for nucleic acids (see Chapter 4). The structure of ATP is similar to those of other nucleotides, but two things

about ATP make it especially useful to cells. First, ATP releases a relatively large amount of energy when hydrolyzed to ADP and P_i. Second, ATP can phosphorylate (donate a phosphate group to) many different molecules, which gain some of the energy that was stored in the ATP. We will examine these two properties in the discussion that follows.

ATP hydrolysis releases energy

An ATP molecule consists of the nitrogenous base adenine bonded to ribose (a sugar), which is attached to a sequence of three phosphate groups (Figure 8.5A). The hydrolysis of a molecule of ATP yields free energy, as well as ADP and the inorganic phosphate ion (P_i). Thus:



(B) *Luciola cruciata*

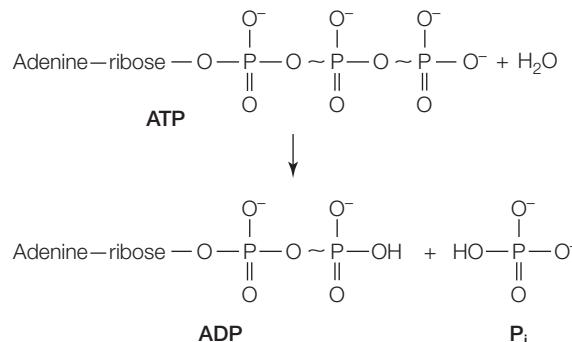


8.5 ATP (A) ATP is richer in energy than its relatives ADP and AMP. (B) Fireflies use ATP to initiate the oxidation of luciferin. This process converts chemical energy into light energy, emitting rhythmic flashes that signal the insect's readiness to mate.

The important property of this reaction is that it is exergonic, releasing free energy. Under standard laboratory conditions, the change in free energy for this reaction (ΔG) is about -7.3 kcal/mol (-30 kJ/mol). However, under cellular conditions, the value can be as much as -14 kcal/mol. We give both values here because you will encounter both values, and you should be aware of their origins. Both are correct, but in different conditions.

Two characteristics of ATP account for the free energy released by the loss of one or two of its phosphate groups:

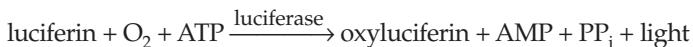
- The free energy of the P—O bond between phosphate groups (called a phosphoric acid anhydride bond) is much higher than the energy of the O—H bond that forms after hydrolysis. So some usable energy is released by hydrolysis.



- Because phosphate groups are negatively charged and so repel each other, it takes energy to get phosphates near enough to each other to make the covalent bond that links them together (e.g., to add a phosphate to ADP to make ATP). Some of this energy is conserved when the third phosphate is attached.

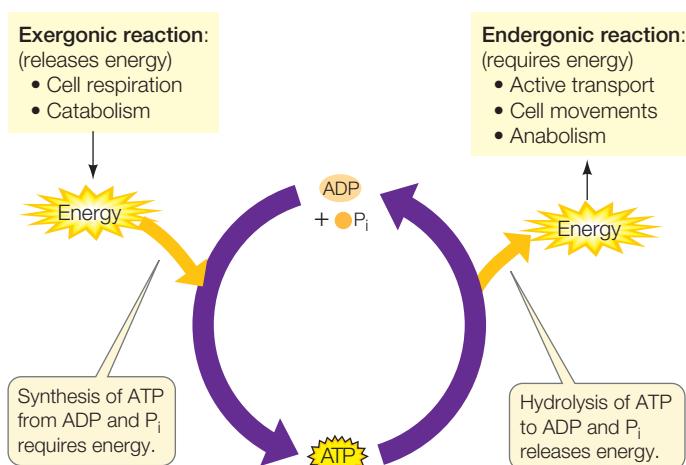
A molecule of ATP can be hydrolyzed either to ADP and P_i, or to adenosine monophosphate (AMP) and a pyrophosphate ion (P₂O₇⁴⁻; commonly abbreviated to PP_i). Cells use the energy released by ATP hydrolysis to fuel endergonic reactions (such as the biosynthesis of complex molecules), for active transport, and for movement. Another interesting example of the use of ATP involves converting its chemical energy into light energy.

BIOLUMINESCENCE The production of light by living organisms is referred to as **bioluminescence** (Figure 8.5B). It is an example of an endergonic reaction driven by ATP hydrolysis that involves an interconversion of energy forms (chemical to light). The chemical that becomes luminescent is called luciferin (after the light-bearing fallen angel, Lucifer):



This reaction and the enzyme that catalyzes it (luciferase) occur in a wide variety of organisms in addition to the familiar firefly. These include a variety of marine organisms, microorganisms, worms, and mushrooms. The light is generally used to avoid predators or for signaling to mates.

Soft-drink companies use the firefly proteins luciferin and luciferase to detect bacterial contamination. Where there are

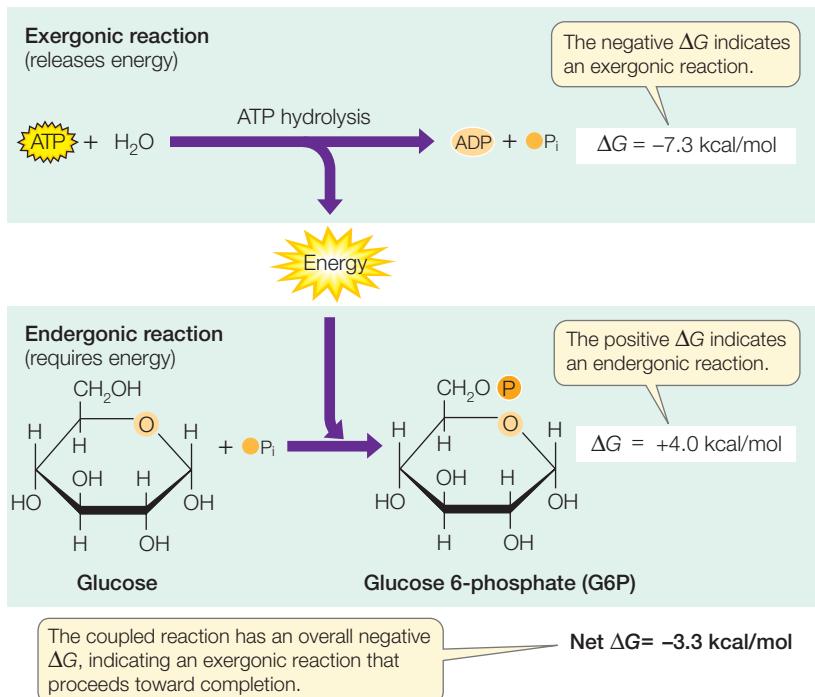


8.6 Coupling of Reactions Exergonic cellular reactions release the energy needed to make ATP from ADP. The energy released from the conversion of ATP back to ADP can be used to fuel endergonic reactions.

yourBioPortal.com

GO TO Web Activity 8.1 • ATP and Coupled Reactions

living cells there is ATP, and when the firefly proteins encounter ATP and oxygen, they give off light. Thus, a sample of soda that lights up in the test is contaminated with bacteria and is discarded.



8.7 Coupling of ATP Hydrolysis to an Endergonic Reaction The addition of phosphate derived from the hydrolysis of ATP to glucose forms the molecule glucose 6-phosphate (in a reaction catalyzed by hexokinase). ATP hydrolysis is exergonic and the energy released drives the second reaction, which is endergonic.

ATP couples exergonic and endergonic reactions

As we have just seen, the hydrolysis of ATP is exergonic and yields ADP, P_i , and free energy. The reverse reaction, the formation of ATP from ADP and P_i , is endergonic and consumes as much free energy as is released by the hydrolysis of ATP:



Many different exergonic reactions in the cell can provide the energy to convert ADP into ATP. For eukaryotes and many prokaryotes, the most important of these reactions is cellular respiration, in which some of the energy released from fuel molecules is captured in ATP. The formation and hydrolysis of ATP constitute what might be called an “energy-coupling cycle,” in which ADP picks up energy from exergonic reactions to become ATP, which then donates energy to endergonic reactions. ATP is the common component of these reactions and is the agent of coupling, as illustrated in **Figure 8.6**.

Coupling of exergonic and endergonic reactions is very common in metabolism. Free energy is captured and retained in the P—O bonds of ATP. ATP then diffuses to another site in the cell, where its hydrolysis releases the free energy to drive an endergonic reaction. For example, the formation of glucose 6-phosphate from glucose (**Figure 8.7**), which has a positive ΔG (is endergonic), will not proceed without the input of free energy from ATP hydrolysis, which has a negative ΔG (is exergonic). The overall ΔG for the coupled reactions (when the two ΔG s are added together) is negative. Hence the reactions proceed exergonically when they are coupled, and glucose 6-phosphate is synthesized. As you will see in Chapter 9, this is the initial reaction in the catabolism of glucose.

An active cell requires the production of millions of molecules of ATP per second to drive its biochemical machinery. An ATP molecule is typically consumed within a second of its formation. At rest, an average person produces and hydrolyzes about 40 kg of ATP per day—as much as some people weigh. This means that each ATP molecule undergoes about 10,000 cycles of synthesis and hydrolysis every day!

8.2 RECAP

ATP is the “energy currency” of cells. Some of the free energy released by exergonic reactions can be captured in the form of ATP. This energy can then be released by ATP hydrolysis and used to drive endergonic reactions.

- How does ATP store energy? See p. 153
- What are coupled reactions? See p. 155 and Figure 8.7

ATP is synthesized and used up very rapidly. But these biochemical reactions could not proceed so rapidly without the help of enzymes.

8.3 What Are Enzymes?

When we know the change in free energy (ΔG) of a reaction, we know where the equilibrium point of the reaction lies: the more negative ΔG is, the further the reaction proceeds toward completion. However, ΔG tells us nothing about the *rate* of a reaction—the speed at which it moves toward equilibrium. The reactions that cells depend on have spontaneous rates that are so slow that the cells would not survive without a way to speed up the reactions. That is the role of catalysts: substances that speed up reactions without themselves being permanently altered. A catalyst does not cause a reaction to occur that would not proceed without it, but merely increases the rate of the reaction, allowing equilibrium to be approached more rapidly. This is an important point: *no catalyst makes a reaction occur that cannot otherwise occur*.

Most biological catalysts are proteins called *enzymes*. Although we will focus here on proteins, some catalysts—perhaps the earliest ones in the origin of life—are RNA molecules called ribozymes (see Section 4.3). A biological catalyst, whether protein or RNA, is a framework or scaffold within which chemical catalysis takes place. This molecular framework binds the reactants and can participate in the reaction itself; however, such participation does not permanently change the enzyme. The catalyst ends up in exactly the same chemical condition after a reaction as before it. Over time, cells have evolved to utilize proteins rather than RNA as catalysts in most biochemical reactions, probably because of the great diversity in the three-dimensional structures of proteins, and because of the variety of chemical functions provided by their functional groups (see Figure 3.1).

In this section we will discuss the energy barrier that controls the rate of a chemical reaction. Then we will focus on the roles of enzymes: how they interact with specific reactants, how they lower the energy barrier, and how they permit reactions to proceed more quickly.

To speed up a reaction, an energy barrier must be overcome

An exergonic reaction may release a great deal of free energy, but take place very slowly. Such reactions are slow because there is an energy barrier between reactants and products. Think about the propane stove we describe in Section 2.3.

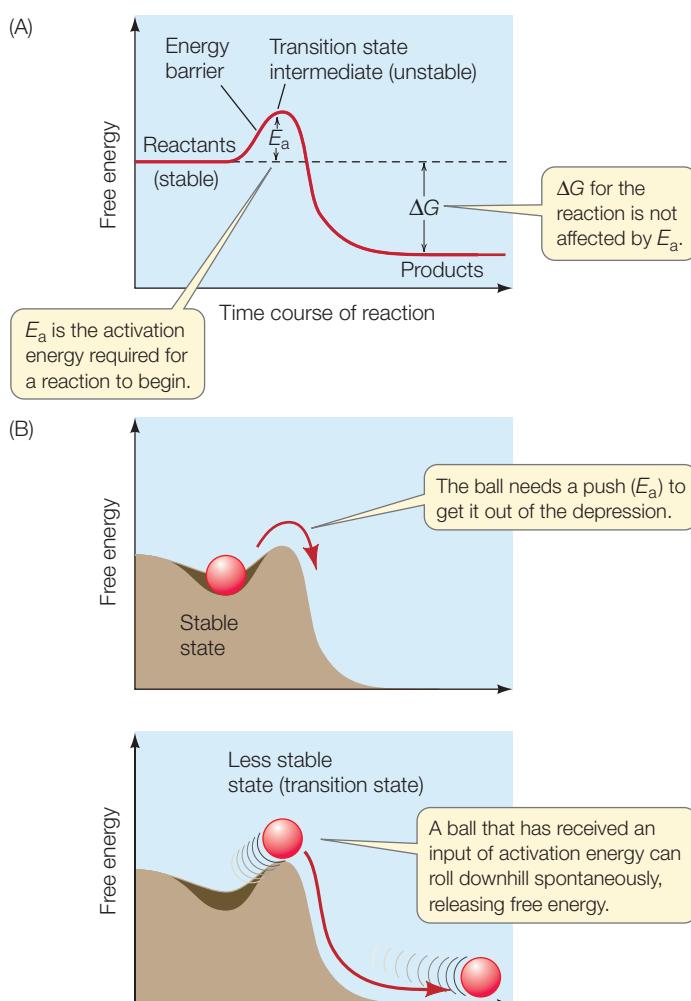
The burning of propane ($C_3H_8 + 5 O_2 \rightarrow 3 CO_2 + 4 H_2O + \text{energy}$) is an exergonic reaction—energy is released in the form of heat and light. Once started, the reaction goes to completion: all of the propane reacts with oxygen to form carbon dioxide and water vapor.

Because burning propane liberates so much energy, you might expect this reaction to proceed rapidly whenever propane is exposed to oxygen. But this does not happen; propane will start burning only if a spark, an input of energy such as a burning match, is provided. A spark is needed because there is an energy barrier between the reactants and the products.

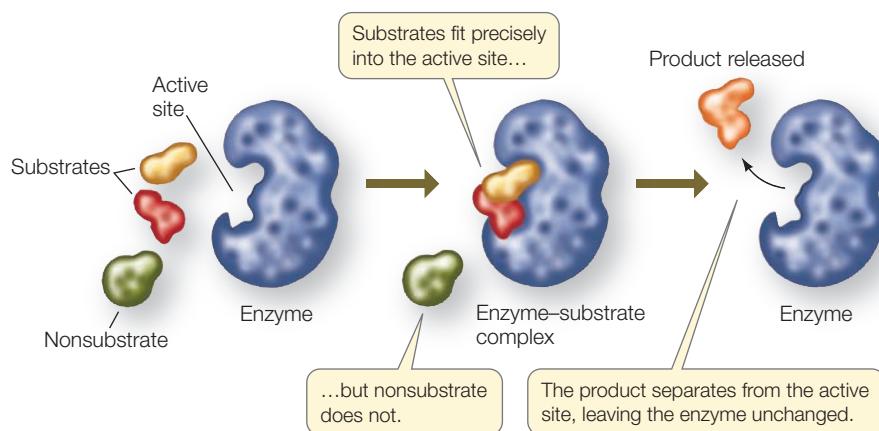
In general, exergonic reactions proceed only after the reactants are pushed over the energy barrier by some added energy.

The energy barrier thus represents the amount of energy needed to start the reaction, known as the **activation energy** (E_a) (Figure 8.8A). Recall the ball rolling down the hill in Figure 8.3. The ball has a lot of potential energy at the top of the hill. However, if it is stuck in a small depression, it will not roll down the hill, even though that action is exergonic. To start the ball rolling, a small amount of energy (activation energy) is needed to push it out of the depression (Figure 8.8B). In a chemical reaction, the activation energy is the energy needed to change the reactants into unstable molecular forms called transition-state intermediates.

Transition-state intermediates have higher free energies than either the reactants or the products. Their bonds may be stretched and therefore unstable. Although the amount of activation energy needed for different reactions varies, it is often small compared with the change in free energy of the reaction. The activation energy put in to start a reaction is recovered during the ensuing “downhill” phase of the reaction, so it is not a part of the net free energy change, ΔG (see Figure 8.8A).



8.8 Activation Energy Initiates Reactions (A) In any chemical reaction, an initial stable state must become less stable before change is possible. (B) A ball on a hillside provides a physical analogy to the biochemical principle graphed in (A).



Where does the activation energy come from? In any collection of reactants at room or body temperatures, the molecules are moving around. A few are moving fast enough that their kinetic energy can overcome the energy barrier, enter the transition state, and react. However, the reaction takes place very slowly at room or body temperatures. If the system were heated, all the reactant molecules would move faster and have more kinetic energy, and the reaction would speed up. You have probably used this technique in the chemistry laboratory.

However, adding enough heat to increase the average kinetic energy of the molecules would not work in living systems. Such a nonspecific approach would accelerate all reactions, including destructive ones such as the denaturation of proteins (see Figure 3.9). A more effective way to speed up a reaction in a living system is to lower the energy barrier by bringing the reactants close together. In living cells, enzymes and ribozymes accomplish this task.

Enzymes bind specific reactants at their active sites

Catalysts increase the rates of chemical reactions. Most nonbiological catalysts are nonspecific. For example, powdered platinum catalyzes virtually any reaction in which molecular hydrogen (H_2) is a reactant. In contrast, most biological catalysts are highly specific. An enzyme or ribozyme usually recognizes and binds to only one or a few closely related reactants, and it catalyzes only a single chemical reaction. In the discussion that follows, we focus on enzymes, but remember that similar rules of chemical behavior apply to ribozymes as well.

In an enzyme-catalyzed reaction, the reactants are called **substrates**. Substrate molecules bind to a particular site on the enzyme, called the active site, where catalysis takes place (Figure 8.9). The specificity of an enzyme results from the exact three-dimensional shape and structure of its active site, into which only a narrow range of substrates can fit. Other molecules—with different shapes, different functional groups, and different properties—cannot fit properly and bind to the active site. This specificity is comparable to the specific binding of a membrane transport protein or receptor protein to its specific ligand, as described in Chapters 6 and 7.

The names of enzymes reflect their functions and often end with the suffix “ase.” For example the enzyme lactase, which you encountered in the opening story for this chapter, catalyzes the hy-

8.9 Enzyme and Substrate An enzyme is a protein catalyst with an active site capable of binding one or more substrate molecules.

drolysis of lactose but not another disaccharide, sucrose. The enzyme hexokinase accelerates the phosphorylation of glucose, but not ribose, to make glucose 6-phosphate (see Figure 8.7).

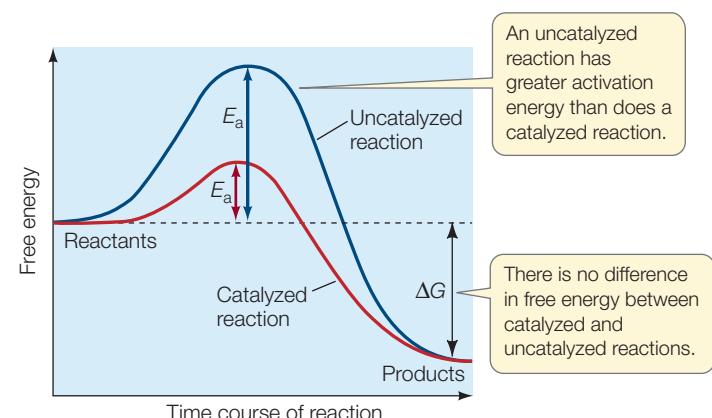
The binding of a substrate to the active site of an enzyme produces an **enzyme–substrate complex (ES)** that is held together by one or more means, such as hydrogen bonding, electrical attraction, or temporary covalent bonding. The enzyme–substrate complex gives rise to product and free enzyme:



where E is the enzyme, S is the substrate, P is the product, and ES is the enzyme–substrate complex. The free enzyme (E) is in the same chemical form at the end of the reaction as at the beginning. While bound to the substrate, it may change chemically, but by the end of the reaction it has been restored to its initial form and is ready to bind more substrate.

Enzymes lower the energy barrier but do not affect equilibrium

When reactants are bound to the enzyme, forming an enzyme–substrate complex, they require less activation energy than the transition-state species of the corresponding uncatalyzed reaction (Figure 8.10). Thus the enzyme lowers the energy barrier for the reaction—it offers the reaction an easier path, speeding it up. When an enzyme lowers the energy bar-



8.10 Enzymes Lower the Energy Barrier Although the activation energy is lower in an enzyme-catalyzed reaction than in an uncatalyzed reaction, the energy released is the same with or without catalysis. In other words, E_a is lower, but ΔG is unchanged. A lower activation energy means the reaction will take place at a faster rate.

rier, both the forward and the reverse reactions speed up, so the enzyme-catalyzed overall reaction proceeds toward equilibrium more rapidly than the uncatalyzed reaction. *The final equilibrium is the same with or without the enzyme.* Similarly, adding an enzyme to a reaction does not change the difference in free energy (ΔG) between the reactants and the products (see Figure 8.10).

Enzymes can change the rate of a reaction substantially. For example, if 600 molecules of a protein with arginine as its terminal amino acid just sit in solution, the protein molecules tend toward disorder and the terminal peptide bonds break, releasing the arginines (ΔS increases). Without an enzyme this is a very slow reaction—it takes about 7 years for half (300) of the proteins to undergo this reaction. However, with the enzyme carboxypeptidase A catalyzing the reaction, the 300 arginines are released in half a second! The important consequence of this for living cells is not difficult to imagine. Such speeds make new realities possible.

8.3 RECAP

A chemical reaction requires a “push” over the energy barrier to get started. Enzymes provide this activation energy by binding specific reactants (substrates).

- Explain how the structure of an enzyme makes that enzyme specific. **See p. 157 and Figure 8.9**
- What is the relationship between an enzyme and the equilibrium point of a reaction? **See pp. 157–158**

Now that you have a general understanding of the structures, functions, and specificities of enzymes, let’s see how they work to speed up chemical reactions between the substrate molecules.

8.4 How Do Enzymes Work?

During and after the formation of the enzyme–substrate complex, chemical interactions occur. These interactions contribute directly to the breaking of old bonds and the formation of new ones. In catalyzing a reaction, an enzyme may use one or more mechanisms.

Enzymes can orient substrates

When free in solution, substrates are moving from place to place randomly while at the same time vibrating, rotating, and tumbling around. They may not have the proper orientation to interact when they collide. Part of the activation energy needed to start a reaction is used to bring together specific atoms so that bonds can form (Figure 8.11A). For example, if acetyl coenzyme A (acetyl CoA) and oxaloacetate are to form citrate (a step in the metabolism of glucose; see Section 9.2), the two substrates must be oriented so that the carbon atom of the methyl group of acetyl CoA can form a covalent bond with the carbon atom of the carbonyl group of oxaloacetate. The active site of the enzyme

citrate synthase has just the right shape to bind these two molecules so that these atoms are adjacent.

Enzymes can induce strain in the substrate

Once a substrate has bound to its active site, an enzyme can cause bonds in the substrate to stretch, putting it in an unstable transition state (Figure 8.11B). For example, lysozyme is a protective enzyme abundant in tears and saliva that destroys invading bacteria by cleaving polysaccharide chains in their cell walls. Lysozyme’s active site “stretches” the bonds of the bacterial polysaccharide, rendering the bonds unstable and more reactive to lysozyme’s other substrate, water.

Enzymes can temporarily add chemical groups to substrates

The side chains (R groups) of an enzyme’s amino acids may be direct participants in making its substrates more chemically reactive (Figure 8.11C).

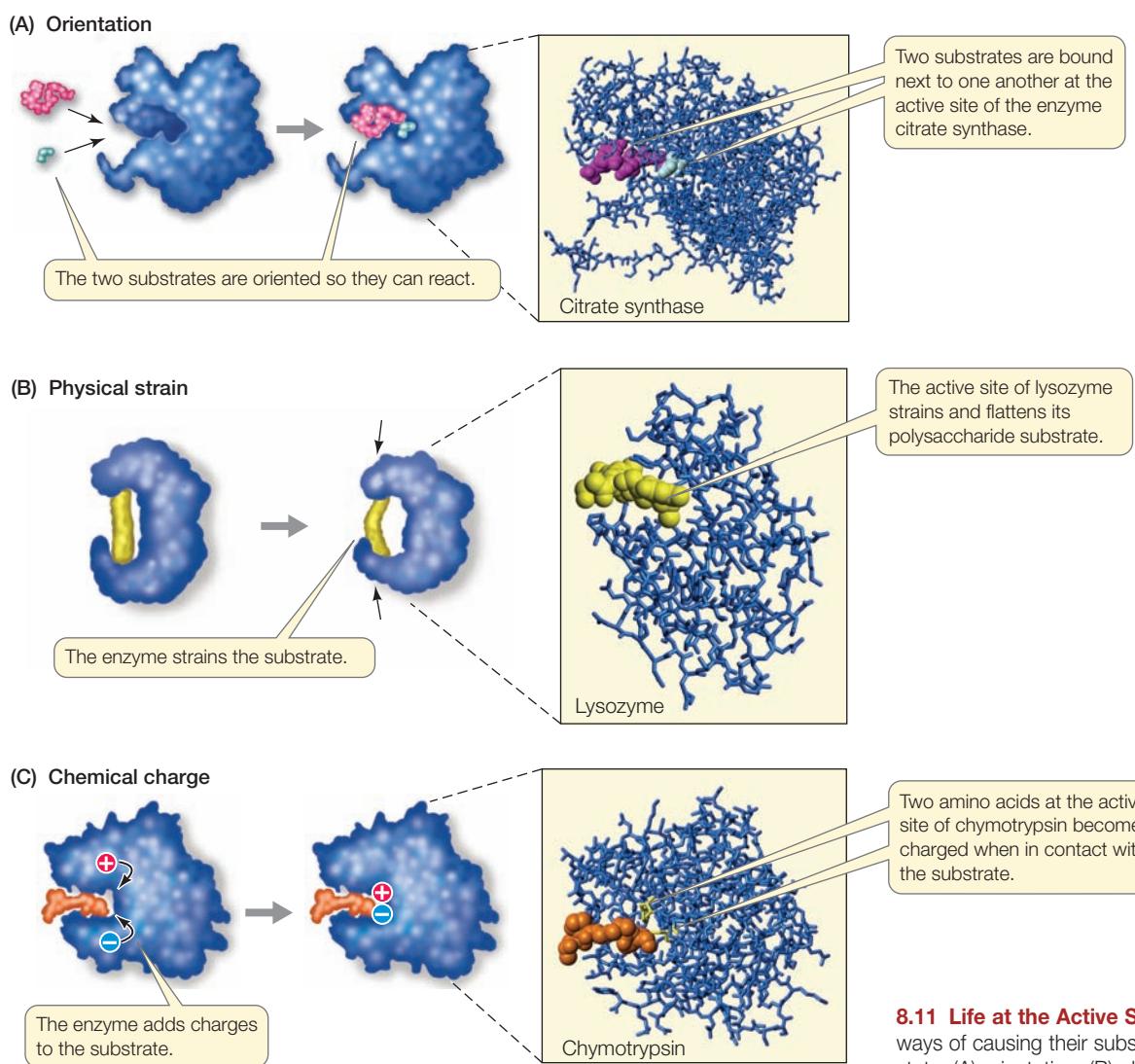
- In *acid–base catalysis*, the acidic or basic side chains of the amino acids in the active site transfer H⁺ to or from the substrate, destabilizing a covalent bond in the substrate, and permitting it to break.
- In *covalent catalysis*, a functional group in a side chain forms a temporary covalent bond with a portion of the substrate.
- In *metal ion catalysis*, metal ions such as copper, iron, and manganese, which are often firmly bound to side chains of enzymes, can lose or gain electrons without detaching from the enzymes. This ability makes them important participants in oxidation–reduction reactions, which involve the loss or gain of electrons.

Molecular structure determines enzyme function

Most enzymes are much larger than their substrates. An enzyme is typically a protein containing hundreds of amino acids and may consist of a single folded polypeptide chain or of several subunits (see Section 3.2). Its substrate is generally a small molecule or a small part of a large molecule. The active site of the enzyme is usually quite small, not more than 6–12 amino acids. Two questions arise from these observations:

- What features of the active site allow it to recognize and bind the substrate?
- What is the role of the rest of the huge protein?

THE ACTIVE SITE IS SPECIFIC TO THE SUBSTRATE The remarkable ability of an enzyme to select exactly the right substrate depends on a precise interlocking of molecular shapes and interactions of chemical groups at the active site. The binding of the substrate to the active site depends on the same kinds of forces that maintain the tertiary structure of the enzyme: hydrogen bonds, the attraction and repulsion of electrically charged groups, and hydrophobic interactions.



8.11 Life at the Active Site Enzymes have several ways of causing their substrates to enter the transition state: (A) orientation, (B) physical strain, and (C) chemical charge.

In 1894, the German chemist Emil Fischer compared the fit between an enzyme and its substrate to that of a lock and key. Fischer's model persisted for more than half a century with only indirect evidence to support it. The first direct evidence came in 1965, when David Phillips and his colleagues at the Royal Institution in London crystallized the enzyme lysozyme and determined its tertiary structure using the technique of X-ray crystallography (described in Section 13.2). They observed a pocket in lysozyme that neatly fits its substrate (see Figure 8.11B).

AN ENZYME CHANGES SHAPE WHEN IT BINDS A SUBSTRATE Just as a membrane receptor protein may undergo precise changes in conformation upon binding to its ligand (see Chapter 7), some enzymes change their shapes when they bind their substrate(s). These shape changes, which are called **induced fit**, expose the active site (or sites) of the enzyme.

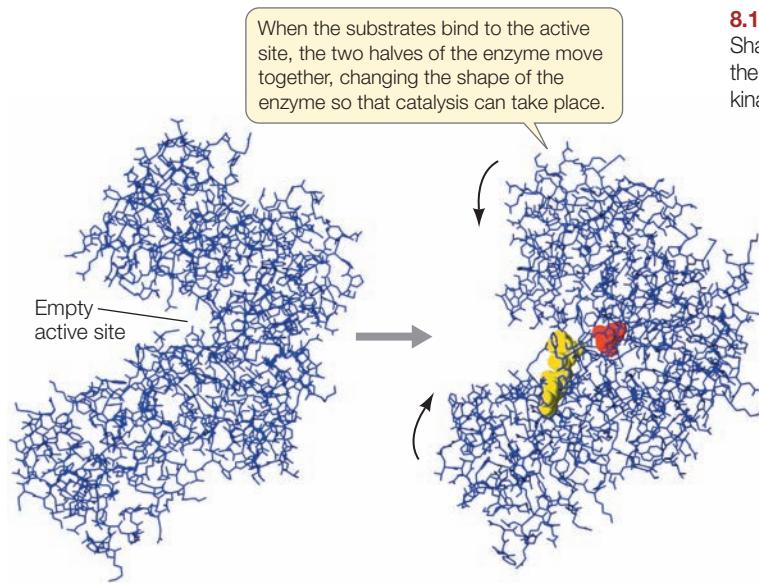
An example of induced fit can be seen in the enzyme hexokinase (see Figure 8.7), which catalyzes the reaction



Induced fit brings reactive side chains from the hexokinase active site into alignment with the substrates (Figure 8.12), facilitating its catalytic mechanisms. Equally important, the folding of hexokinase to fit around the substrates (glucose and ATP) excludes water from the active site. This is essential, because if water were present, the ATP could be hydrolyzed to ADP and P_i. But since water is absent, the transfer of a phosphate from ATP to glucose is favored.

Induced fit at least partly explains why enzymes are so large. The rest of the macromolecule may have three roles:

- It provides a framework so that the amino acids of the active site are properly positioned in relation to the substrate(s).
- It participates in significant changes in protein shape and structure that result in induced fit.
- It provides binding sites for regulatory molecules (see Section 8.5).



Some enzymes require other molecules in order to function

As large and complex as enzymes are, many of them require the presence of nonprotein chemical “partners” in order to function (**Table 8.1**):

- **Prosthetic groups** are distinctive, non-amino acid atoms or molecular groupings that are permanently bound to their enzymes. An example is a flavin nucleotide, which binds to succinate dehydrogenase, an important enzyme in cellular respiration (see Section 9.2).
- **Cofactors** are inorganic ions such as copper, zinc, and iron that bind to certain enzymes. For example, the cofactor zinc binds to the enzyme alcohol dehydrogenase.

TABLE 8.1
Some Examples of Nonprotein “Partners” of Enzymes

TYPE OF MOLECULE	ROLE IN CATALYZED REACTIONS
COFACTORS	
Iron (Fe^{2+} or Fe^{3+})	Oxidation/reduction
Copper (Cu^+ or Cu^{2+})	Oxidation/reduction
Zinc (Zn^{2+})	Helps bind NAD
COENZYMES	
Biotin	Carries $-\text{COO}^-$
Coenzyme A	Carries $-\text{CO}-\text{CH}_3$
NAD	Carries electrons
FAD	Carries electrons
ATP	Provides/extracts energy
PROSTHETIC GROUPS	
Heme	Binds ions, O_2 , and electrons; contains iron cofactor
Flavin	Binds electrons
Retinal	Converts light energy

8.12 Some Enzymes Change Shape When Substrate Binds to Them

Shape changes result in an induced fit between enzyme and substrate, improving the catalytic ability of the enzyme. Induced fit can be observed in the enzyme hexokinase, seen here with and without its substrates, glucose (red) and ATP (yellow).

- A **coenzyme** is a carbon-containing molecule that is required for the action of one or more enzymes. It is usually relatively small compared with the enzyme to which it temporarily binds.

A coenzyme moves from enzyme to enzyme, adding or removing chemical groups from the substrate. A coenzyme is like a substrate in that it does not permanently bind to the enzyme; it binds to the active site, changes chemically during the reaction, and then separates from the enzyme to participate in other reactions. ATP and ADP, as energy carriers, can be considered coenzymes, even though they are really substrates. The term coenzyme was coined before the functions of these molecules were fully understood. Biochemists continue to use the term, and to be consistent with the field, we will use the term in this book.

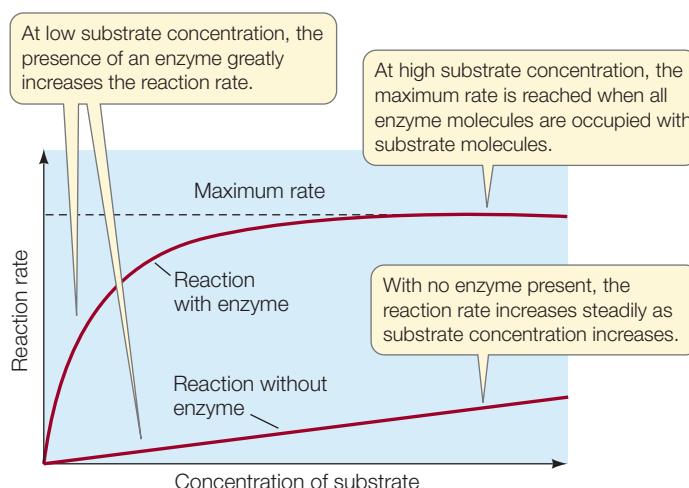
In the next chapter we will encounter other coenzymes that function in energy-harvesting reactions by accepting or donating electrons or hydrogen atoms. In animals, some coenzymes are produced from vitamins—substances that must be obtained from food because they cannot be synthesized by the body. For example, the B vitamin niacin is used to make the coenzyme nicotinamide adenine dinucleotide (NAD).

The substrate concentration affects the reaction rate

For a reaction of the type $\text{A} \rightarrow \text{B}$, the rate of the uncatalyzed reaction is directly proportional to the concentration of A. The higher the concentration of substrate, the more reactions per unit of time. Addition of the appropriate enzyme speeds up the reaction, of course, but it also changes the shape of a plot of rate versus substrate concentration (**Figure 8.13**). For a given concentration of enzyme, the rate of the enzyme-catalyzed reaction initially increases as the substrate concentration increases from zero, but then it levels off. At some point, further increases in the substrate concentration do not significantly increase the reaction rate—the maximum rate has been reached.

Since the concentration of an enzyme is usually much lower than that of its substrate and does not change as substrate concentration changes, what we see is a saturation phenomenon like the one that occurs in facilitated diffusion (see Figure 6.14). When all the enzyme molecules are bound to substrate molecules, the enzyme is working as fast as it can—at its maximum rate. Nothing is gained by adding more substrate, because no free enzyme molecules are left to act as catalysts. Under these conditions the active sites are said to be saturated.

The maximum rate of a catalyzed reaction can be used to measure how efficient the enzyme is—that is, how many molecules of substrate are converted into product per unit of time when there is an excess of substrate present. This *turnover number* ranges from one molecule every two seconds for lysozyme to an amazing 40 million molecules per second for the liver enzyme catalase.



8.13 Catalyzed Reactions Reach a Maximum Rate Because there is usually less enzyme than substrate present, the reaction rate levels off when the enzyme becomes saturated.

8.4 RECAP

Enzymes orient their substrates to bring together specific atoms so that bonds can form. An enzyme can participate in the reaction it catalyzes by temporarily changing shape or destabilizing the enzyme-substrate complex. Some enzymes require cofactors, coenzymes, or prosthetic groups in order to function.

- What are three mechanisms of enzyme catalysis? See p. 158 and Figure 8.11
- What are the chemical roles of coenzymes in enzymatic reactions? See p. 160

Now that you understand more about how enzymes function, let's see how different enzymes work together in a complex organism.

8.5 How Are Enzyme Activities Regulated?

A major characteristic of life is homeostasis—the maintenance of stable internal conditions (see Chapter 40). How does a cell maintain a relatively constant internal environment while thousands of chemical reactions are going on? These chemical reactions operate within *metabolic pathways* in which the product of one reaction is a reactant for the next. The pathway for the metabolism of lactose begins with lactase (as we described in the chapter's opening story), and is just one of many pathways that regulate the internal environment of the cell. These pathways have such diverse functions as the catabolism of glucose to yield energy, CO_2 , and H_2O , and the anabolism of amino acids to yield proteins. Metabolic pathways do not exist in isolation, but interact extensively, and each reaction in each pathway is catalyzed by a specific enzyme.

Within a cell or organism, the presence and activity of enzymes determine the “flow” of chemicals through different metabolic pathways. The amount of enzyme activity, in turn, is controlled in part via the regulation of gene expression. Many signal transduction pathways (described in Chapter 7) end with changes in gene expression, and often the genes that are switched on or off encode enzymes. But the simple presence of an enzyme does not ensure that it is functioning. Another means by which cells can control which pathways are active at a particular time is by the activation or inactivation of enzymes. If one enzyme in the pathway is inactive, that step and all subsequent steps shut down. Thus, enzymes are target points for the regulation of entire sequences of chemical reactions.

Regulation of the rates at which thousands of different enzymes operate contributes to homeostasis within an organism. Such control permits cells to make orderly changes in their functions in response to changes in the external environment. In Chapter 7 we describe a number of enzymes that become activated in signal transduction pathways, illustrating how enzyme activation can dramatically alter cell functions. (For example, see the activation of glycogen phosphorylase in Figure 7.20.)

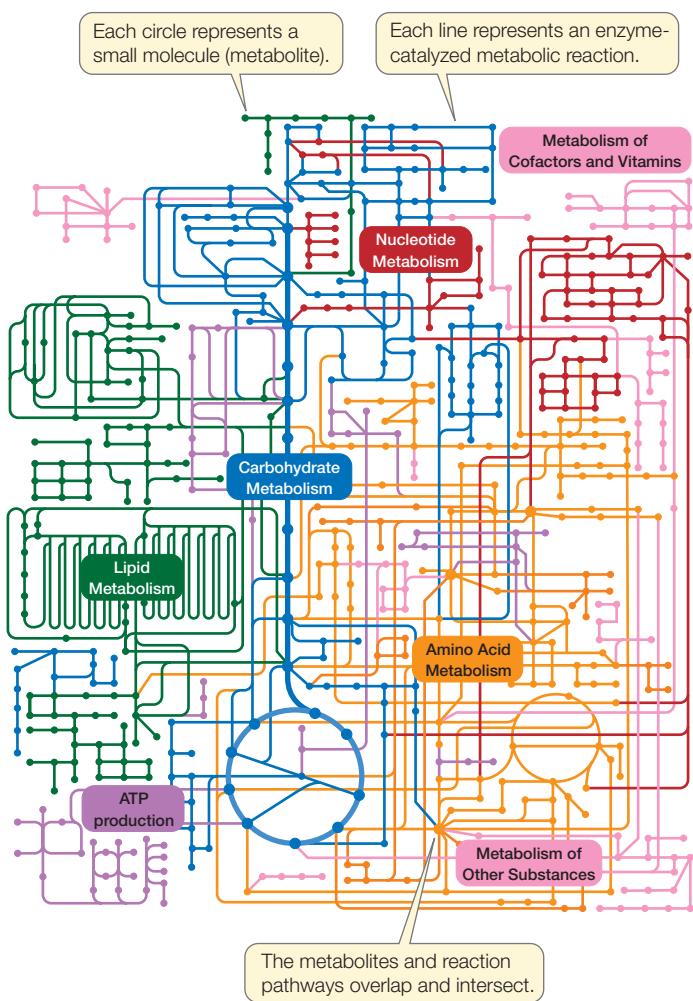
The flow of chemicals such as carbon atoms through interacting metabolic pathways can be studied, but this process becomes complicated quickly, because each pathway influences the others. Computer algorithms are used to model these pathways and show how they mesh in an interdependent system (Figure 8.14). Such models can help predict what will happen if the concentration of one molecule or another is altered. This new field of biology is called **systems biology**, and it has numerous applications.

In this section we will investigate the roles of enzymes in organizing and regulating metabolic pathways. In living cells, enzymes can be activated or inhibited in various ways, and there are also mechanisms for controlling the rates at which some enzymes catalyze reactions. We will also examine how the environment—particularly temperature and pH—affects enzyme activity.

Enzymes can be regulated by inhibitors

Various chemical inhibitors can bind to enzymes, slowing down the rates of the reactions they catalyze. Some inhibitors occur naturally in cells; others are artificial. Naturally occurring inhibitors regulate metabolism; artificial ones can be used to treat disease, to kill pests, or to study how enzymes work. In some cases the inhibitor binds the enzyme irreversibly, and the enzyme becomes permanently inactivated. In other cases the inhibitor has reversible effects; it can separate from the enzyme, allowing the enzyme to function fully as before. The removal of a natural reversible inhibitor increases an enzyme's rate of catalysis.

IRREVERSIBLE INHIBITION If an inhibitor covalently binds to certain side chains at the active site of an enzyme, it will permanently inactivate the enzyme by destroying its capacity to interact with its normal substrate. An example of an irreversible inhibitor is DIPF (diisopropyl phosphorofluoridate), which



8.14 Metabolic Pathways The complex interactions of metabolic pathways can be modeled by the tools of systems biology. In cells, the main elements controlling these pathways are enzymes.

reacts with serine (Figure 8.15). DIPF is an irreversible inhibitor of acetylcholinesterase, whose operation is essential for the normal functioning of the nervous system. Because of their effect on acetylcholinesterase, DIPF and other similar compounds are classified as nerve gases, and were developed for biological warfare. One of these compounds, Sarin, was used in an attack on the Tokyo subway in 1995, resulting in a dozen deaths and the hospitalization of hundreds more. The widely used insecticide malathion is a derivative of DIPF that inhibits only insect acetylcholinesterase, not the mammalian enzyme. The irreversible inhibition of enzymes is of practical use to humans, but this form of regulation is not common in the cell, because the enzyme is permanently inactivated and cannot be recycled. Instead, cells use reversible inhibition.

REVERSIBLE INHIBITION In some cases an inhibitor is similar enough to a particular enzyme's natural substrate to bind non-covalently to its active site, yet different enough that the enzyme catalyzes no chemical reaction. While such a molecule is bound to the enzyme, the natural substrate cannot enter the active site

and the enzyme is unable to function. Such a molecule is called a **competitive inhibitor** because it competes with the natural substrate for the active site (Figure 8.16A). In this case, the inhibition is reversible. When the concentration of the competitive inhibitor is reduced, it detaches from the active site, and the enzyme is active again.

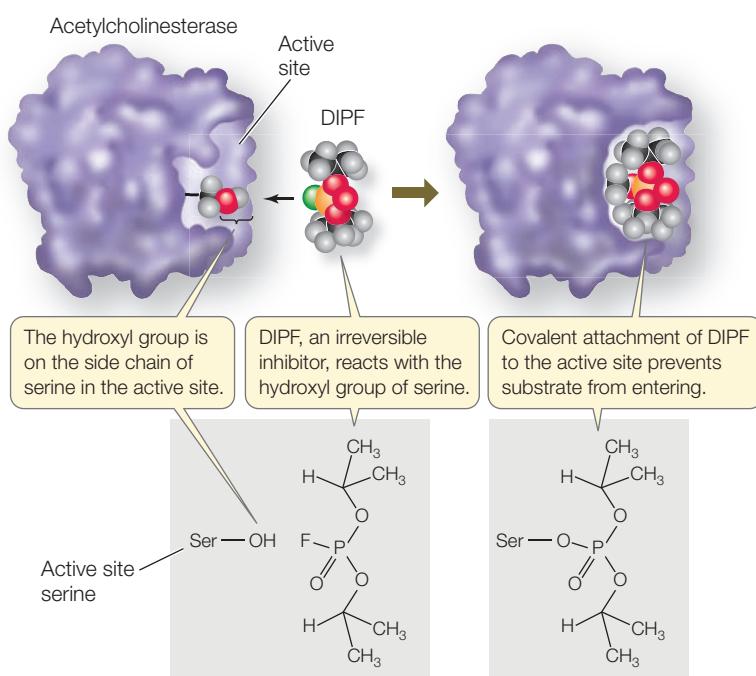
A **noncompetitive inhibitor** binds to an enzyme at a site distinct from the active site. This binding causes a change in the shape of the enzyme that alters its activity (Figure 8.16B). The active site may no longer bind the substrate, or if it does, the rate of product formation may be reduced. Like competitive inhibitors, noncompetitive inhibitors can become unbound, so their effects are reversible.

Allosteric enzymes control their activity by changing shape

The change in enzyme shape due to noncompetitive inhibitor binding is an example of allostery (*allo*, “different”; *stereos*, “shape”). **Allosteric regulation** occurs when an effector molecule binds to a site other than the active site of an enzyme, *inducing the enzyme to change its shape*. The change in shape alters the affinity of the active site for the substrate, and so the rate of the reaction is changed.

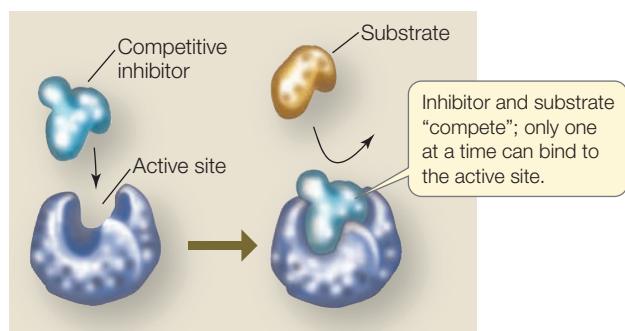
Often, an enzyme will exist in the cell in more than one possible shape (Figure 8.17):

- The *active form* of the enzyme has the proper shape for substrate binding.
- The *inactive form* of the enzyme has a shape that cannot bind the substrate.

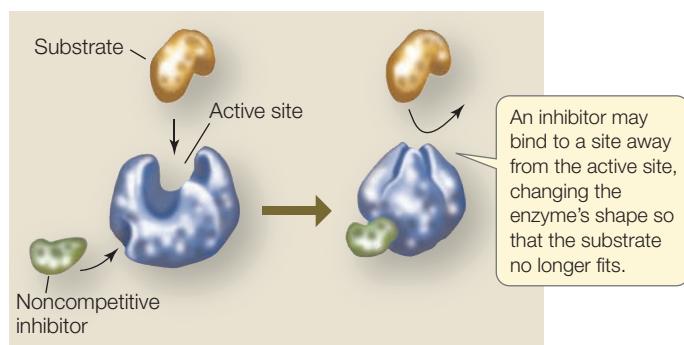


8.15 Irreversible Inhibition DIPF forms a stable covalent bond with the side chain of the amino acid serine at the active site of the enzyme acetylcholinesterase, thus irreversibly disabling the enzyme.

(A) Competitive inhibition



(B) Noncompetitive inhibition



8.16 Reversible Inhibition (A) A competitive inhibitor binds temporarily to the active site of an enzyme. (B) A noncompetitive inhibitor binds temporarily to the enzyme at a site away from the active site. In both cases, the enzyme's function is disabled for only as long as the inhibitor remains bound.

yourBioPortal.com
GO TO Animated Tutorial 8.1 • Enzyme Catalysis

Other molecules, collectively referred to as effectors, can influence which form the enzyme takes:

- Binding of an inhibitor to a site separate from the active site can stabilize the inactive form of the enzyme, making it less likely to convert to the active form.
- The active form can be stabilized by the binding of an activator to another site on the enzyme.

Like substrate binding, the binding of inhibitors and activators to their regulatory sites (also called allosteric sites) is highly specific. Most (but not all) enzymes that are allosterically regulated are proteins with quaternary structure; that is, they are made up of multiple polypeptide subunits. The polypeptide that has the active site is called the catalytic subunit. The allosteric sites are often on different polypeptides, called the regulatory subunits.

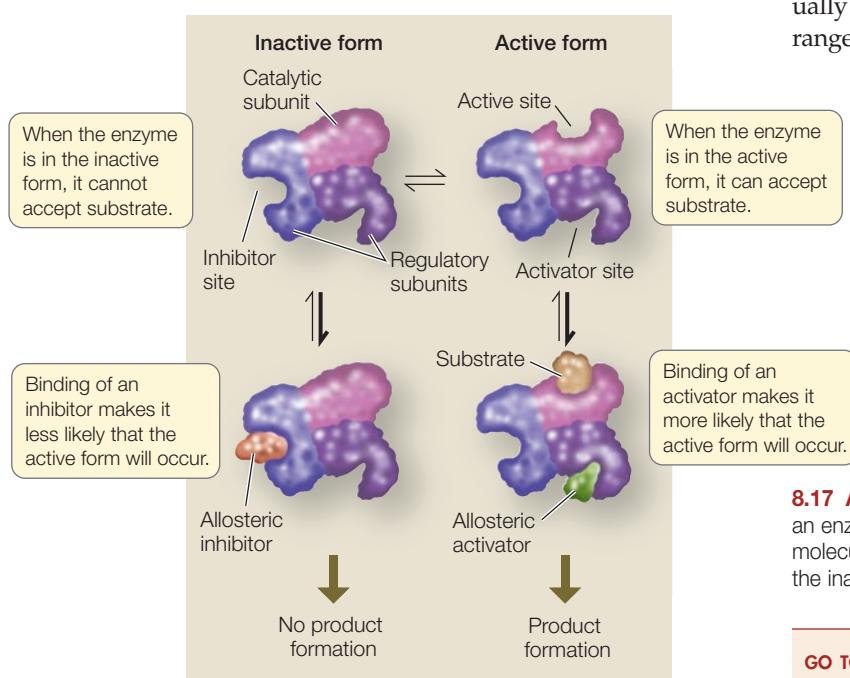
Some enzymes have multiple subunits containing active sites, and the binding of substrate to one of the active sites causes allosteric effects. When substrate binds to one subunit, there is a slight change in protein structure that influences the adjacent subunit. The slight change to the second subunit makes its active site more likely to bind to the substrate. So the reaction speeds up as the sites become sequentially activated.

As a result, an allosteric enzyme with multiple active sites and a nonallosteric enzyme with a single active site differ greatly in their reaction rates when the substrate concentration is low. Graphs of reaction rates plotted against substrate concentrations show this relationship. For a nonallosteric enzyme, the plot looks like that in **Figure 8.18A**. The reaction rate first increases sharply with increasing substrate concentration, then tapers off to a constant maximum rate as the supply of enzyme becomes saturated.

The plot for a multisubunit allosteric enzyme is radically different, having a sigmoid (S-shaped) appearance (**Figure 8.18B**). At low substrate concentrations, the reaction rate increases gradually as substrate concentration increases. But within a certain range, the reaction rate is extremely sensitive to relatively small changes in substrate concentration. In addition, allosteric enzymes are very sensitive to low concentrations of inhibitors. Because of this sensitivity, allosteric enzymes are important in regulating entire metabolic pathways.

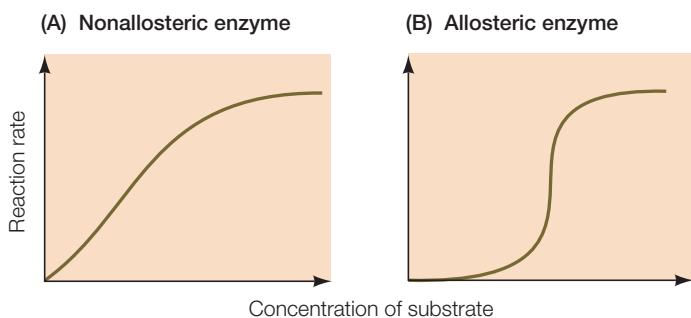
Allosteric effects regulate metabolism

Metabolic pathways typically involve a starting material, various intermediate products, and an end product that is used for some purpose by the cell. In each pathway there are a number of reactions, each



8.17 Allosteric Regulation of Enzymes Active and inactive forms of an enzyme can be interconverted, depending on the binding of effector molecules at sites other than the active site. Binding an inhibitor stabilizes the inactive form and binding an activator stabilizes the active form.

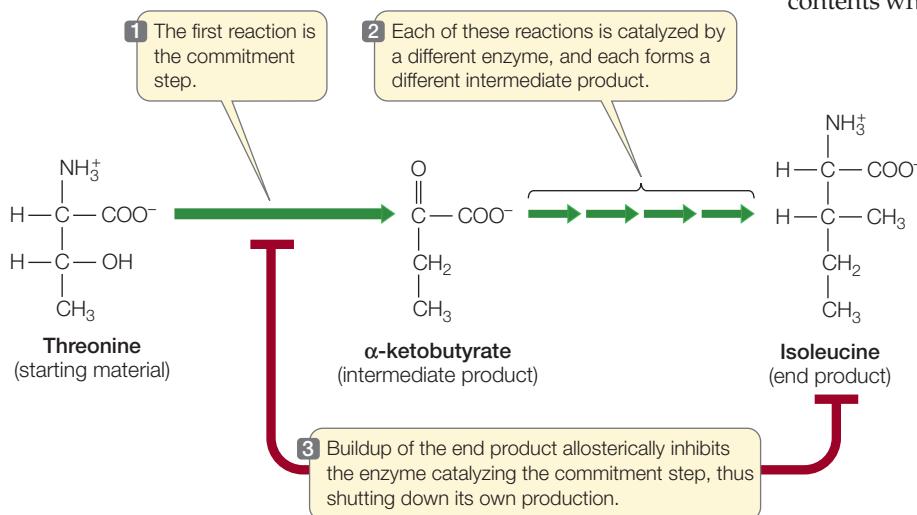
yourBioPortal.com
GO TO Animated Tutorial 8.2 • Allosteric Regulation of Enzymes



8.18 Allostery and Reaction Rate The number of active sites on an enzyme determines how the rate of the enzyme-catalyzed reaction changes as substrate concentration increases. A sigmoid curve (B) is typical for an enzyme with multiple subunits, each with an active site. After one subunit binds the substrate, changes in structure make it more likely that the next subunit will also bind substrate. So the reaction speeds up more rapidly than in the case of an enzyme with a single active site (A).

forming an intermediate product and each catalyzed by a different enzyme. The first step in a pathway is called the commitment step, meaning that once this enzyme-catalyzed reaction occurs, the “ball is rolling,” and the other reactions happen in sequence, leading to the end product. But what if the cell has no need for that product—for example, if that product is available from its environment in adequate amounts? It would be energetically wasteful for the cell to continue making something it does not need.

One way to avoid this problem is to shut down the metabolic pathway by having the final product inhibit the enzyme that catalyzes the commitment step (**Figure 8.19**). Often this inhibition occurs allosterically. When the end product is present at a high concentration, some of it binds to an allosteric site on the commitment step enzyme, thereby causing it to become inactive. Thus, the final product acts as a *noncompetitive inhibitor* (described earlier in this section) of the first enzyme in the pathway. This mechanism is known as feedback inhibition or end-product inhibition. We will describe many other examples of such inhibition in later chapters.



Enzymes are affected by their environment

Enzymes enable cells to perform chemical reactions and carry out complex processes rapidly without using the extremes of temperature and pH employed by chemists in the laboratory. However, because of their three-dimensional structures and the chemistry of the side chains in their active sites, enzymes (and their substrates) are highly sensitive to changes in temperature and pH. In Section 3.2 we describe the general effects of these environmental factors on proteins. Here we will examine their effects on enzyme function (which, of course, depends on enzyme structure and chemistry).

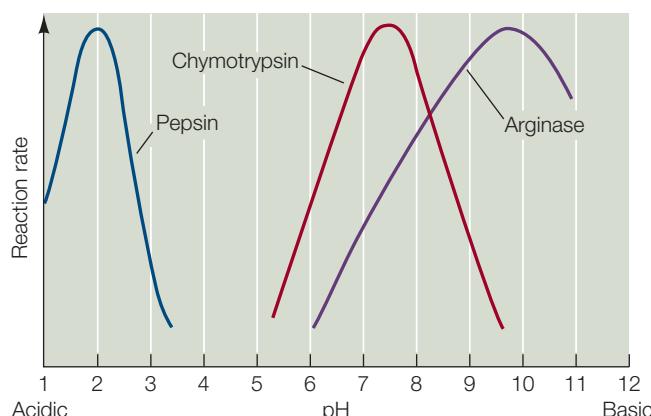
pH AFFECTS ENZYME ACTIVITY The rates of most enzyme-catalyzed reactions depend on the pH of the solution in which they occur. While the water inside cells is generally at a neutral pH of 7, the presence of acids, bases, and buffers can alter this. Each enzyme is most active at a particular pH; its activity decreases as the solution is made more acidic or more basic than the ideal (optimal) pH (**Figure 8.20**). As an example, consider the human digestive system (see Section 51.3). The pH inside the human stomach is highly acidic, around pH 1.5. Many enzymes that hydrolyze macromolecules, such as proteases, have pH optima in the neutral range. So when food enters the small intestine, a buffer (bicarbonate) is secreted into the intestine to raise the pH to 6.5. This allows the hydrolytic enzymes to be active and digest the food.

Several factors contribute to this effect. One factor is ionization of the carboxyl, amino, and other groups on either the substrate or the enzyme. In neutral or basic solutions, carboxyl groups ($-\text{COOH}$) release H^+ to become negatively charged carboxylate groups ($-\text{COO}^-$). On the other hand, in neutral or acidic solutions, amino groups ($-\text{NH}_2$) accept H^+ to become positively charged $-\text{NH}_3^+$ groups (see the discussion of acids and bases in Section 2.4). Thus, in a neutral solution, an amino group is electrically attracted to a carboxyl group on another molecule or another part of the same molecule, because both groups are ionized and have opposite charges. If the pH changes, however, the ionization of these groups may change. For example, at a low pH (high H^+ concentration, such as the stomach contents where the enzyme pepsin is active), the excess H^+ may

react with $-\text{COO}^-$ to form $-\text{COOH}$. If this happens, the group is no longer charged and cannot interact with other charged groups in the protein, so the folding of the protein may be altered. If such a change occurs at the active site of an enzyme, the enzyme may no longer be able to bind to its substrate.

8.19 Feedback Inhibition of Metabolic Pathways

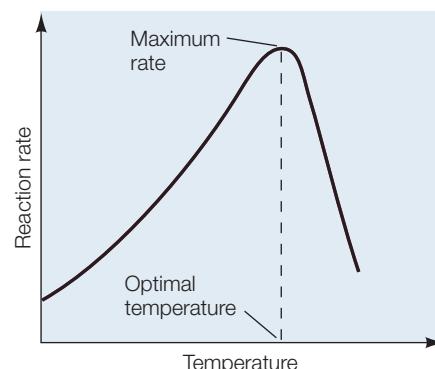
The first reaction in a metabolic pathway is referred to as the commitment step. It is often catalyzed by an enzyme that can be allosterically inhibited by the end product of the pathway. The specific pathway shown here is the synthesis of isoleucine from threonine in bacteria. It is typical of many enzyme-catalyzed biosynthetic pathways.



8.20 pH Affects Enzyme Activity An enzyme catalyzes its reaction at a maximum rate. The activity curve for each enzyme peaks at its optimal pH. For example, pepsin is active in the acidic environment of the stomach, while chymotrypsin is active in the small intestine.

TEMPERATURE AFFECTS ENZYME ACTIVITY In general, warming increases the rate of a chemical reaction because a greater proportion of the reactant molecules have enough kinetic energy to provide the activation energy for the reaction. Enzyme-catalyzed reactions are no different (Figure 8.21). However, temperatures that are too high inactivate enzymes, because at high temperatures enzyme molecules vibrate and twist so rapidly that some of their noncovalent bonds break. When an enzyme's tertiary structure is changed by heat it loses its function. Some enzymes denature at temperatures only slightly above that of the human body, but a few are stable even at the boiling point (or freezing point) of water. All enzymes, however, have an optimal temperature for activity.

Individual organisms adapt to changes in the environment in many ways, one of which is based on groups of enzymes, called *isozymes*, that catalyze the same reaction but have different chemical compositions and physical properties. Different isozymes within a given group may have different optimal temperatures. The rainbow trout, for example, has several isozymes of the enzyme acetylcholinesterase. If a rainbow trout is transferred from warm water to near-freezing water (2°C), the fish produces an isozyme of acetylcholinesterase that is different from the one it produces at the higher temperature. The new isozyme has a lower optimal temperature, allowing the fish's nervous system to perform normally in the colder water.



8.21 Temperature Affects Enzyme Activity Each enzyme is most active at a particular optimal temperature. At higher temperatures the enzyme becomes denatured and inactive; this explains why the activity curve falls off abruptly at temperatures above the optimal.

In general, enzymes adapted to warm temperatures do not denature at those temperatures because their tertiary structures are held together largely by covalent bonds, such as charge interactions or disulfide bridges, instead of the more heat-sensitive weak chemical interactions. Most enzymes in humans are more stable at high temperatures than those of the bacteria that infect us, so that a moderate fever tends to denature bacterial enzymes, but not our own.

8.5 RECAP

The rates of most enzyme-catalyzed reactions are affected by interacting molecules (such as inhibitors and activators) and by environmental factors (such as temperature and pH).

- What is the difference between reversible and irreversible enzyme inhibition? See pp. 161–162
- How are allosteric enzymes regulated? See pp. 162–163 and Figure 8.17
- Explain the concept of feedback inhibition. How might the reactions shown in Figure 8.19 fit into a systems diagram such as the one shown in Figure 8.14?

CHAPTER SUMMARY

8.1 What Physical Principles Underlie Biological Energy Transformations?

- Energy is the capacity to do work. In a biological system, the usable energy is called **free energy (G)**. The unusable energy is **entropy**, a measure of the disorder in the system.
- Potential energy** is the energy of state or position; it includes the energy stored in chemical bonds. **Kinetic energy** is the energy of motion; it is the type of energy that can do work.

- The **laws of thermodynamics** apply to living organisms. The first law states that energy cannot be created or destroyed. The second law states that energy transformations decrease the amount of energy available to do work (free energy) and increase disorder. Review Figure 8.2
- The **change in free energy (ΔG)** of a reaction determines its point of **chemical equilibrium**, at which the forward and reverse reactions proceed at the same rate.

- An **exergonic reaction** releases free energy and has a negative ΔG . An **endergonic reaction** consumes or requires free energy and has a positive ΔG . Endergonic reactions proceed only if free energy is provided. [Review Figure 8.3](#)
- Metabolism** is the sum of all the biochemical (metabolic) reactions in an organism. **Catabolic reactions** are associated with the breakdown of complex molecules and release energy (are exergonic). **Anabolic reactions** build complexity in the cell and are endergonic.

8.2 What Is the Role of ATP in Biochemical Energetics?

- Adenosine triphosphate (ATP)** serves as an energy currency in cells. Hydrolysis of ATP releases a relatively large amount of free energy.
- The **ATP cycle** couples exergonic and endergonic reactions, harvesting free energy from exergonic reactions, and providing free energy for endergonic reactions. [Review Figure 8.6](#), [WEB ACTIVITY 8.1](#)

8.3 What Are Enzymes?

- The rate of a chemical reaction is independent of ΔG , but is determined by the **energy barrier**. **Enzymes** are protein catalysts that affect the rates of biological reactions by lowering the energy barrier, supplying the **activation energy (E_a)** needed to initiate reactions. [Review Figure 8.10](#), [WEB ACTIVITY 8.2](#)
- A **substrate** binds to the enzyme's **active site**—the site of catalysis—forming an **enzyme–substrate complex**. Enzymes are highly specific for their substrates.

8.4 How Do Enzymes Work?

- At the active site, a substrate can be oriented correctly, chemically modified, or strained. As a result, the substrate readily forms its **transition state**, and the reaction proceeds. [Review Figure 8.11](#)
- Binding substrate causes many enzymes to change shape, exposing their active site(s) and allowing catalysis. The change in enzyme shape caused by substrate binding is known as **induced fit**. [Review Figure 8.12](#)
- Some enzymes require other substances, known as **cofactors**, to carry out catalysis. **Prosthetic groups** are permanently bound to enzymes; **coenzymes** are not. A coenzyme can be considered a substrate, as it is changed by the reaction and then released from the enzyme.
- Substrate concentration affects the rate of an enzyme-catalyzed reaction.

8.5 How Are Enzyme Activities Regulated?

- Metabolism is organized into pathways in which the product of one reaction is a reactant for the next reaction. Each reaction in the pathway is catalyzed by an enzyme.
- Enzyme activity is subject to regulation. Some inhibitors bind irreversibly to enzymes. Others bind reversibly. [Review Figures 8.15 and 8.16](#), [ANIMATED TUTORIAL 8.1](#)
- An **allosteric effector** binds to a site other than the active site and stabilizes the active or inactive form of an enzyme. [Review Figure 8.17](#), [ANIMATED TUTORIAL 8.2](#)
- The end product of a metabolic pathway may inhibit an enzyme that catalyzes the **commitment step** of that pathway. [Review Figure 8.19](#)
- Enzymes are sensitive to their environments. Both pH and temperature affect enzyme activity. [Review Figures 8.20 and 8.21](#)

SELF-QUIZ

- Coenzymes differ from enzymes in that coenzymes are
 - only active outside the cell.
 - polymers of amino acids.
 - smaller molecules, such as vitamins.
 - specific for one reaction.
 - always carriers of high-energy phosphate.
- Which statement about thermodynamics is true?
 - Free energy is used up in an exergonic reaction.
 - Free energy cannot be used to do work.
 - The total amount of energy can change after a chemical transformation.
 - Free energy can be kinetic but not potential energy.
 - Entropy has a tendency to increase.
- In a chemical reaction,
 - the rate depends on the value of ΔG .
 - the rate depends on the activation energy.
 - the entropy change depends on the activation energy.
 - the activation energy depends on the value of ΔG .
 - the change in free energy depends on the activation energy.
- Which statement about enzymes is *not* true?
 - They usually consist of proteins.
 - They change the rate of the catalyzed reaction.
 - They change the ΔG of the reaction.
 - They are sensitive to heat.
 - They are sensitive to pH.
- The active site of an enzyme
 - never changes shape.
 - forms no chemical bonds with substrates.
 - determines, by its structure, the specificity of the enzyme.
 - looks like a lump projecting from the surface of the enzyme.
 - changes the ΔG of the reaction.
- The molecule ATP is
 - a component of most proteins.
 - high in energy because of the presence of adenine.
 - required for many energy-transforming biochemical reactions.
 - a catalyst.
 - used in some exergonic reactions to provide energy.
- In an enzyme-catalyzed reaction,
 - a substrate does not change.
 - the rate decreases as substrate concentration increases.
 - the enzyme can be permanently changed.
 - strain may be added to a substrate.
 - the rate is not affected by substrate concentration.

8. Which statement about enzyme inhibitors is *not* true?
 - a. A competitive inhibitor binds the active site of the enzyme.
 - b. An allosteric inhibitor binds a site on the active form of the enzyme.
 - c. A noncompetitive inhibitor binds a site other than the active site.
 - d. Noncompetitive inhibition cannot be completely overcome by the addition of more substrate.
 - e. Competitive inhibition can be completely overcome by the addition of more substrate.
 9. Which statement about the feedback inhibition of enzymes is *not* true?
 - a. It is usually exerted through allosteric effects.
- b. It is directed at the enzyme that catalyzes the commitment step in a metabolic pathway.
 - c. It affects the rate of reaction, not the concentration of enzyme.
 - d. It acts by permanently modifying the active site.
 - e. It is an example of reversible inhibition.
10. Which statement about temperature effects is *not* true?
 - a. Raising the temperature may reduce the activity of an enzyme.
 - b. Raising the temperature may increase the activity of an enzyme.
 - c. Raising the temperature may denature an enzyme.
 - d. Some enzymes are stable at the boiling point of water.
 - e. All enzymes have the same optimal temperature.

FOR DISCUSSION

1. What makes it possible for endergonic reactions to proceed in organisms?
2. Consider two proteins: one is an enzyme dissolved in the cytosol of a cell, the other is an ion channel in its plasma membrane. Contrast the structures of the two proteins, indicating at least two important differences.
3. Plot free energy versus the time course of an endergonic reaction, and the same for an exergonic reaction. Include

the activation energy on both plots. Label E_a and ΔG on both graphs.

4. Consider an enzyme that is subject to allosteric regulation. If a competitive inhibitor (not an allosteric inhibitor) is added to a solution containing such an enzyme, the ratio of enzyme molecules in the active form to those in the inactive form increases. Explain this observation.

ADDITIONAL INVESTIGATION

In humans, hydrogen peroxide (H_2O_2) is a dangerous toxin produced as a by-product of several metabolic pathways. The accumulation of H_2O_2 is prevented by its conversion to harmless H_2O , a reaction catalyzed by the appropriately named enzyme catalase. Air pollutants can inhibit this enzyme and leave indi-

viduals susceptible to tissue damage by H_2O_2 . How would you investigate whether catalase has an allosteric or a nonallosteric mechanism, and whether the pollutants are acting as competitive or noncompetitive inhibitors?

9

Pathways that Harvest Chemical Energy

Of mice and marathons

Like success in your biology course, winning a prestigious marathon comes only after a lot of hard work. Distance runners have more mitochondria in the leg muscles than most of us. The chemical energy stored in the bonds of ATP in those mitochondria is converted into mechanical energy to move the muscles.

There are two types of muscle fibers. Most people have about equal proportions of each type. But in a top marathon racer, 90 percent of the body's muscle is made up of so-called *slow-twitch* fibers. Cells of these fibers have lots of mitochondria and use oxygen to break down fats and carbohydrates, forming ATP. In contrast, the muscles of sprinters are about 80 percent *fast-twitch* fibers, which have fewer mitochondria. Fast-twitch fibers generate short bursts of ATP in the absence of O₂, but the ATP is soon used up. Extensive research with athletes has shown that training can improve the efficiency of blood circulation to the muscle fibers, providing more oxygen, and can even change the ratio of fast-twitch to slow-twitch fibers.

Now enter Marathon Mouse. No, this is not a cartoon character or a computer game, but a very real mouse that was genetically programmed by Ron Evans at the Salk Institute to express high levels of the protein PPARδ in its muscles. This protein is a receptor located inside cell nuclei, where it regulates the transcription of genes involved with the breakdown of fat to yield ATP. Evans's mouse was supposed to break down fats better, and thus be leaner—but there was an unexpected bonus. With high levels of PPARδ came an increase in slow-twitch fibers and a decrease in fast-twitch ones. It was as if the mouse had been in marathon training for a long time!

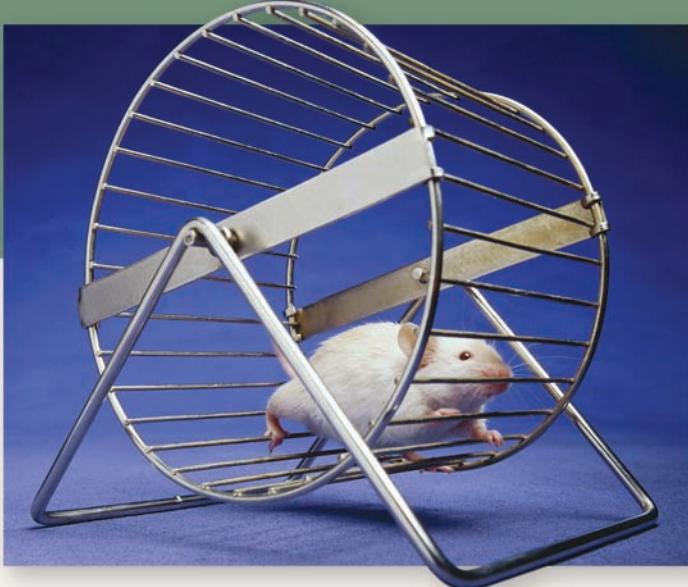
Marathon mice are leaner and meaner than ordinary mice. Leaner, because they are good at burning fat; and meaner in terms of their ability to run long distances. On an exercise wheel, a normal mouse can run for 90 minutes and about a half-mile (900 meters) before it gets tired. PPARδ-enhanced mice can run almost twice as long and twice as far—marks of true distance runners. Could we

also manipulate genes to enhance performance (and fat burning) in humans?

The genetic engineering of people, if it is feasible, is probably far in the future. But implanting genetically altered muscle tissue is actually not such a far-fetched idea, and has already raised concerns over improper athletic enhancement. More likely in the near term is the use of an experimental drug called Aicar, which activates the PPARδ



Marathon Men It takes a lot of training to run a marathon. One of the results of all that training is that the leg muscles become packed with slow-twitch muscle fibers, containing cells rich in energy-metabolizing mitochondria.



Marathon Mouse This mouse can run for much longer than a normal mouse because its energy metabolism has been genetically altered.

protein. When Evans and colleagues gave the drug to normal mice, they achieved the same results as with the genetically modified mice. A test for Aicar in blood and urine has been developed to prevent its use by human athletes to gain a competitive advantage. Of more importance is the drug's potential in the treatment of obesity and diabetes, since the drug stimulates fat breakdown. Obesity is a key part of a disorder called metabolic syndrome, which also includes high blood pressure, heart disease, and diabetes.

The free energy trapped in ATP is the energy you use all the time to fuel both conscious actions, like running a marathon or turning the pages of a book, and your body's automatic actions, such as breathing or contracting your heart muscles.

IN THIS CHAPTER we will describe how cells extract usable energy from food, usually in the form of ATP. We describe the general principles of energy transformations in cells, and illustrate these principles by describing the pathways for the catabolism of glucose in the presence and absence of O₂. Finally, we describe the relationships between the metabolic pathways that use and produce the four biologically important classes of molecules—carbohydrates, fats, proteins, and nucleic acids.

CHAPTER OUTLINE

- 9.1 How Does Glucose Oxidation Release Chemical Energy?
- 9.2 What Are the Aerobic Pathways of Glucose Metabolism?
- 9.3 How Does Oxidative Phosphorylation Form ATP?
- 9.4 How Is Energy Harvested from Glucose in the Absence of Oxygen?
- 9.5 How Are Metabolic Pathways Interrelated and Regulated?

9.1 How Does Glucose Oxidation Release Chemical Energy?

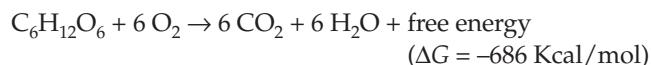
Energy is stored in the covalent bonds of fuels, and it can be released and transformed. Wood burning in a campfire releases energy as heat and light. In cells, fuel molecules release chemical energy that is used to make ATP, which in turn drives endergonic reactions. ATP is central to the energy transformations of all living organisms. Photosynthetic organisms use energy from sunlight to synthesize their own fuels, as we describe in Chapter 10. In nonphotosynthetic organisms, the most common chemical fuel is the sugar glucose (C₆H₁₂O₆). Other molecules, including other carbohydrates, fats, and proteins, can also supply energy. However, to release their energy they must be converted into glucose or intermediate compounds that can enter into the various pathways of glucose metabolism.

In this section we explore how cells obtain energy from glucose by the chemical process of oxidation, which is carried out through a series of metabolic pathways. Five principles govern metabolic pathways:

- A complex chemical transformation occurs in a series of separate reactions that form a metabolic pathway.
- Each reaction is catalyzed by a specific enzyme.
- Most metabolic pathways are similar in all organisms, from bacteria to humans.
- In eukaryotes, many metabolic pathways are compartmentalized, with certain reactions occurring inside specific organelles.
- Each metabolic pathway is regulated by key enzymes that can be inhibited or activated, thereby determining how fast the reactions will go.

Cells trap free energy while metabolizing glucose

As we saw in Section 2.3, the familiar process of combustion (burning) is very similar to the chemical processes that release energy in cells. If glucose is burned in a flame, it reacts with oxygen gas (O₂), forming carbon dioxide and water and releasing energy in the form of heat. The balanced equation for the complete combustion reaction is



This is an oxidation-reduction reaction. Glucose ($C_6H_{12}O_6$) becomes completely oxidized and six molecules of O_2 are reduced to six molecules of water. The energy that is released can be used to do work. The same equation applies to the overall metabolism of glucose in cells. However, in contrast to combustion, the metabolism of glucose is a multistep pathway—each step is catalyzed by an enzyme, and the process is compartmentalized. Unlike combustion, glucose metabolism is tightly regulated and occurs at temperatures compatible with life.

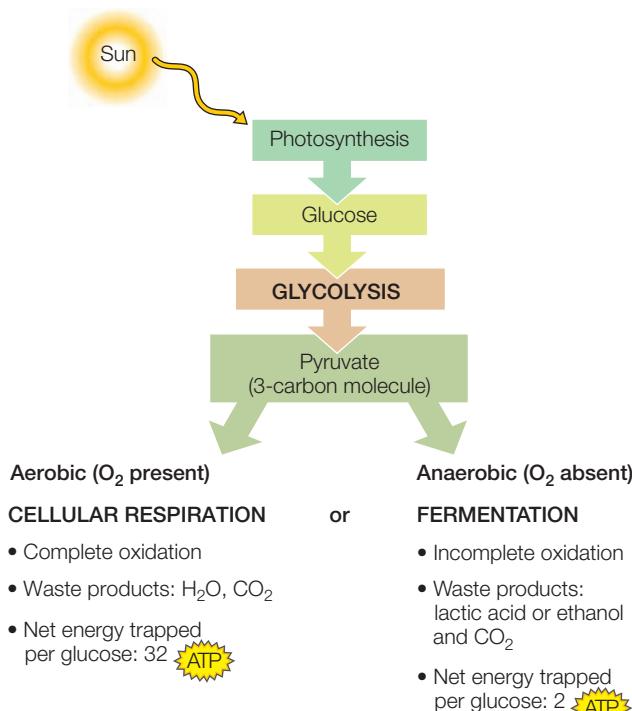
The glucose metabolism pathway “traps” the energy stored in the covalent bonds of glucose and stores it instead in ATP molecules, via the phosphorylation reaction:



As we introduce in Chapter 8, ATP is the energy currency of cells. The energy trapped in ATP can be used to do cellular work—such as movement of muscles or active transport across membranes—just as the energy captured from combustion can be used to do work.

The change in free energy (ΔG) resulting from the complete conversion of glucose and O_2 to CO_2 and water, whether by combustion or by metabolism, is -686 kcal/mol ($-2,870\text{ kJ/mol}$). Thus the overall reaction is highly exergonic and can drive the endergonic formation of a great deal of ATP from ADP and phosphate. Note that in the discussion that follows, “energy” means free energy.

Three metabolic processes harvest the energy in the chemical bonds of glucose: glycolysis, cellular respiration, and fermentation (Figure 9.1). All three processes involve pathways made up of many distinct chemical reactions.



9.1 Energy for Life Living organisms obtain their energy from the food compounds produced by photosynthesis. They convert these compounds into glucose, which they metabolize to trap energy in ATP.

- **Glycolysis** begins glucose metabolism in all cells. Through a series of chemical rearrangements, glucose is converted to two molecules of the three-carbon product **pyruvate**, and a small amount of energy is captured in usable forms. Glycolysis is an **anaerobic** process because it does not require O_2 .

- **Cellular respiration** uses O_2 from the environment, and thus it is **aerobic**. Each pyruvate molecule is completely converted into three molecules of CO_2 through a set of metabolic pathways including pyruvate oxidation, the citric acid cycle, and an electron transport system (the respiratory chain). In the process, a great deal of the energy stored in the covalent bonds of pyruvate is captured to form ATP.

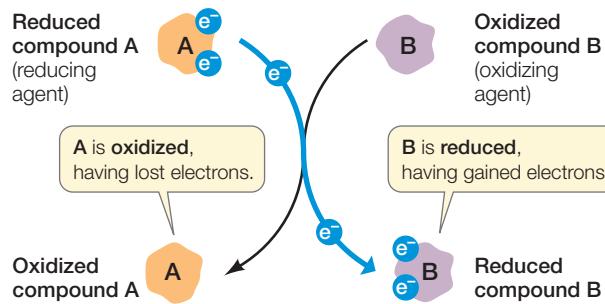
- **Fermentation** does not involve O_2 (it is anaerobic). Fermentation converts pyruvate into lactic acid or ethyl alcohol (ethanol), which are still relatively energy-rich molecules. Because the breakdown of glucose is incomplete, much less energy is released by fermentation than by cellular respiration.

Redox reactions transfer electrons and energy

As is illustrated in Figure 8.6, the addition of a phosphate group to ADP to make ATP is an endergonic reaction that can extract and transfer energy from exergonic to endergonic reactions. Another way of transferring energy is to transfer electrons. A reaction in which one substance transfers one or more electrons to another substance is called an oxidation–reduction reaction, or **redox** reaction.

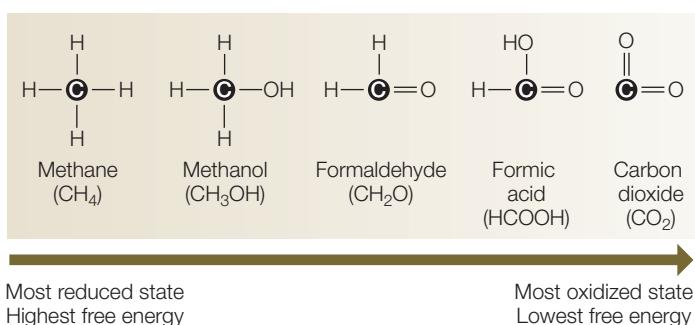
- **Reduction** is the gain of one or more electrons by an atom, ion, or molecule.
- **Oxidation** is the loss of one or more electrons.

Oxidation and reduction *always occur together*: as one chemical is oxidized, the electrons it loses are transferred to another chemical, reducing it. In a redox reaction, we call the reactant that becomes reduced an oxidizing agent and the one that becomes oxidized a reducing agent:



In both the combustion and the metabolism of glucose, glucose is the reducing agent (electron donor) and O_2 is the oxidizing agent (electron acceptor).

Although oxidation and reduction are always defined in terms of traffic in electrons, it is often helpful to think in terms of the gain or loss of hydrogen atoms. Transfers of hydrogen atoms involve transfers of electrons ($H = H^+ + e^-$). So when a molecule loses hydrogen atoms, it becomes oxidized.

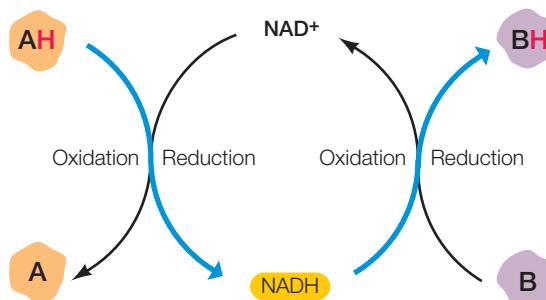


9.2 Oxidation, Reduction, and Energy The more oxidized a carbon atom in a molecule is, the less its free energy.

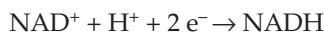
In general, the more reduced a molecule is, the more energy is stored in its covalent bonds (Figure 9.2). In a redox reaction, some energy is transferred from the reducing agent to the reduced product. The rest remains in the reducing agent or is lost to entropy. As we will see, some of the key reactions of glycolysis and cellular respiration are highly exergonic redox reactions.

The coenzyme NAD⁺ is a key electron carrier in redox reactions

Section 8.4 describes the role of coenzymes, small molecules that assist in enzyme-catalyzed reactions. ADP acts as a coenzyme when it picks up energy released in an exergonic reaction and packages it to form ATP. On the other hand, the coenzyme nicotinamide adenine dinucleotide (NAD⁺) acts as an electron carrier in redox reactions:



As you can see, NAD⁺ exists in two chemically distinct forms, one oxidized (NAD⁺) and the other reduced (NADH) (Figure 9.3). Both forms participate in redox reactions. The reduction reaction



is actually the transfer of a proton (the hydrogen ion, H⁺) and two electrons, which are released by the accompanying oxidization reaction.

The electrons do not remain with the coenzyme. Oxygen is highly electronegative and readily accepts electrons from NADH. The oxidation of NADH by O₂ (which occurs in several steps)



is highly exergonic, with a ΔG of -52.4 kcal/mol (-219 kJ/mol). Note that the oxidizing agent appears here as "½ O₂" instead of "O." This notation emphasizes that it is molecular oxygen, O₂, that acts as the oxidizing agent.

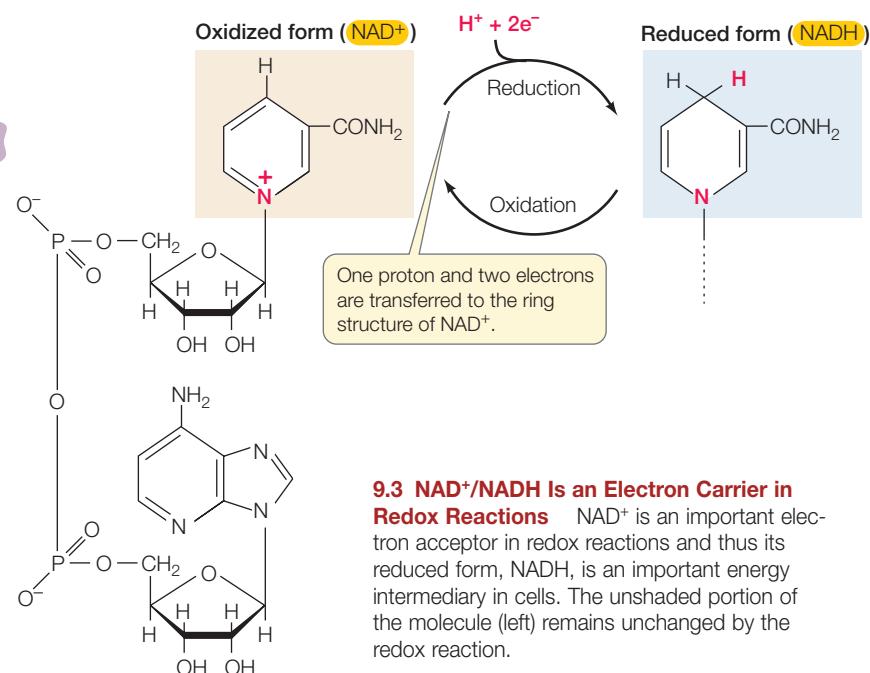
Just as a molecule of ATP can be thought of as a package of about 12 kcal/mol (50 kJ/mol) of free energy, NADH can be thought of as a larger package of free energy (approximately 50 kcal/mol, or 200 kJ/mol). NAD⁺ is a common electron carrier in cells, but not the only one. Another carrier, flavin adenine dinucleotide (FAD), also transfers electrons during glucose metabolism.

An overview: Harvesting energy from glucose

The energy-harvesting processes in cells use different combinations of metabolic pathways depending on the presence or absence of O₂:

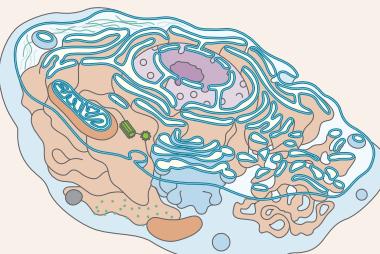
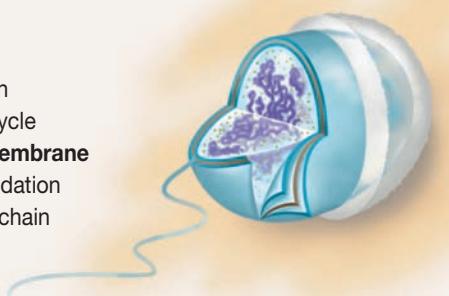
- Under aerobic conditions, when O₂ is available as the final electron acceptor, four pathways operate (Figure 9.4A). Glycolysis is followed by the three pathways of cellular respiration: pyruvate oxidation, the citric acid cycle (also called the Krebs cycle or the tricarboxylic acid cycle), and electron transport/ATP synthesis (also called the respiratory chain).
- Under anaerobic conditions when O₂ is unavailable, pyruvate oxidation, the citric acid cycle, and the respiratory chain do not function, and the pyruvate produced by glycolysis is further metabolized by fermentation (Figure 9.4B).

These five metabolic pathways occur in different locations in the cell (Table 9.1).



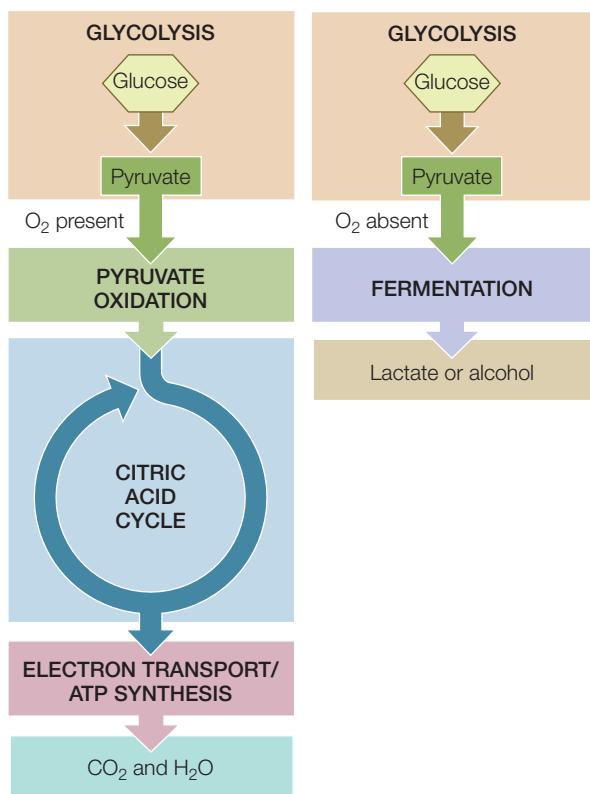
9.3 NAD⁺/NADH Is an Electron Carrier in Redox Reactions NAD⁺ is an important electron acceptor in redox reactions and thus its reduced form, NADH, is an important energy intermediary in cells. The unshaded portion of the molecule (left) remains unchanged by the redox reaction.

TABLE 9.1
Cellular Locations for Energy Pathways in Eukaryotes and Prokaryotes

EUKARYOTES	PROKARYOTES
	
External to mitochondrion Glycolysis Fermentation	In cytoplasm Glycolysis Fermentation
Inside mitochondrion Inner membrane Respiratory chain Matrix Citric acid cycle Pyruvate oxidation	On plasma membrane Pyruvate oxidation Respiratory chain

yourBioPortal.com
GO TO Web Activity 9.1 • Energy Pathways in Cells

- (A) Glycolysis and cellular respiration (B) Glycolysis and fermentation



9.4 Energy-Producing Metabolic Pathways Energy-producing reactions can be grouped into five metabolic pathways: glycolysis, pyruvate oxidation, the citric acid cycle, the respiratory chain/ATP synthesis, and fermentation. (A) The three lower pathways occur only in the presence of O_2 and are collectively referred to as cellular respiration. (B) When O_2 is unavailable, glycolysis is followed by fermentation.

yourBioPortal.com
GO TO Web Activity 9.2 • Glycolysis and Fermentation

9.1 RECAP

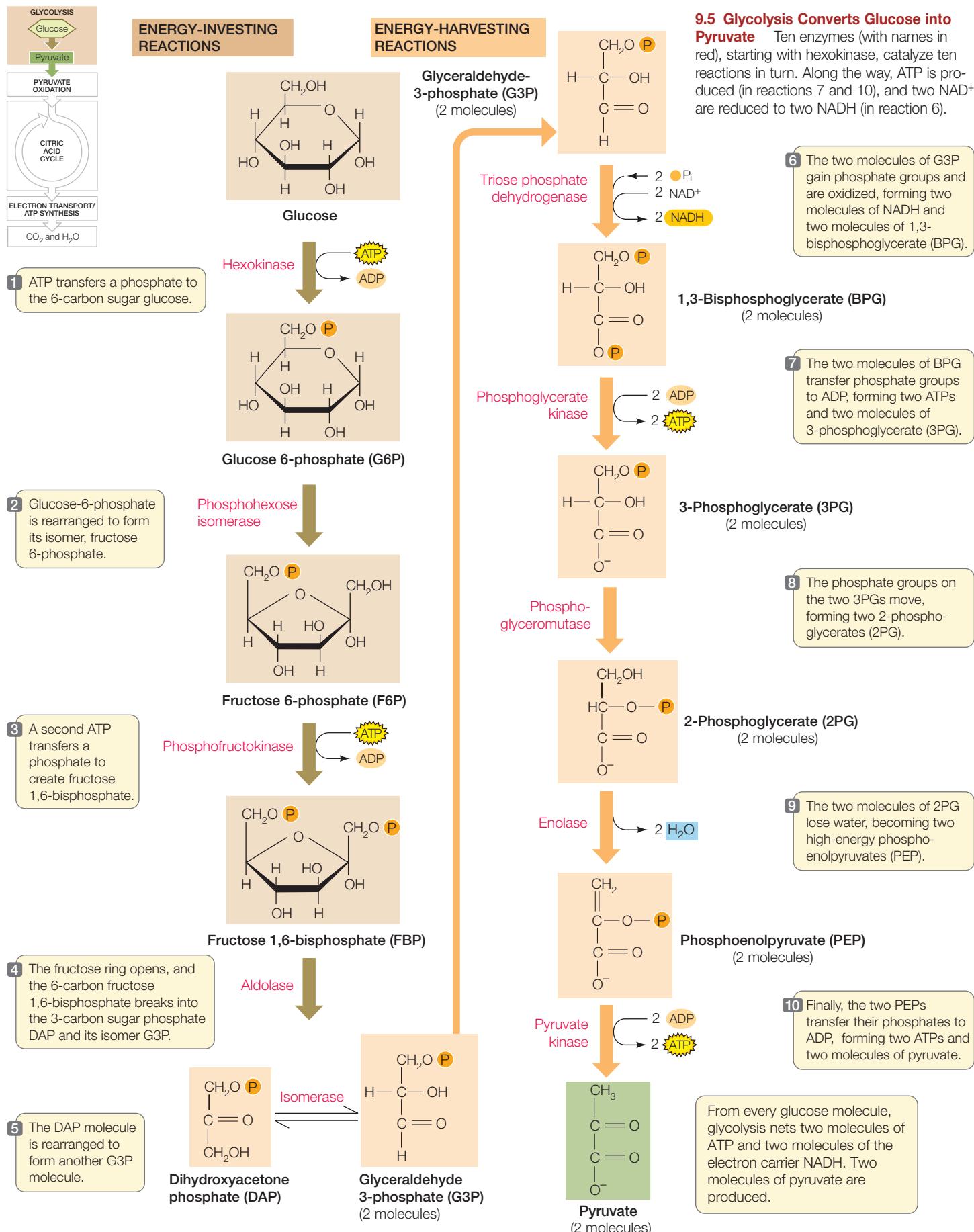
The free energy released from the oxidation of glucose is trapped in the form of ATP. Five metabolic pathways combine in different ways to produce ATP, which supplies the energy for myriad other reactions in living cells.

- What principles govern metabolic pathways in cells? [See p. 169](#)
- Describe how the coupling of oxidation and reduction transfers energy from one molecule to another. [See pp. 170–171](#)
- Explain the roles of NAD^+ and O_2 with respect to electrons in a redox reaction. [See p. 171 and Figure 9.3](#)

Now that you have an overview of the metabolic pathways that harvest energy from glucose, let's take a closer look at the three pathways involved in aerobic catabolism: glycolysis, pyruvate oxidation, and the citric acid cycle.

9.2 What Are the Aerobic Pathways of Glucose Metabolism?

The aerobic pathways of glucose metabolism oxidize glucose completely to CO_2 and H_2O . Initially, the glycolysis reactions convert the six-carbon glucose molecule to two 3-carbon pyruvate molecules (**Figure 9.5**). Pyruvate is then converted to CO_2 in a second series of reactions beginning with pyruvate oxidation and followed by the citric acid cycle. In addition to generating CO_2 , the oxidation events are coupled with the reduction of electron carriers, mostly NAD^+ . So much of the chemical energy in the C—C and C—H bonds of glucose is transferred to NADH. Ultimately, this energy will be transferred to ATP, but this comes in a separate series of reactions involving electron transport, called the respiratory chain. In the respiratory chain, redox reactions result in the oxidative phosphorylation of ADP by ATP synthase. We will begin our consideration of the metabolism of glucose with a closer look at glycolysis.



Glycolysis takes place in the cytosol. It converts glucose into pyruvate, produces a small amount of energy, and generates no CO_2 . During glycolysis, some of the covalent bonds between carbon and hydrogen in the glucose molecule are oxidized, releasing some of the stored energy. The ten enzyme-catalyzed reactions of glycolysis result in the net production of two molecules of pyruvate (pyruvic acid), two molecules of ATP, and two molecules of NADH. Glycolysis can be divided into two stages: energy-investing reactions that consume ATP, and energy-harvesting reactions that produce ATP (see Figure 9.5). We'll begin with the energy-investing reactions.

The energy-investing reactions 1–5 of glycolysis require ATP

Using Figure 9.5 as a guide, let's work our way through the glycolytic pathway.

Two of the reactions (1 and 3 in Figure 9.5), involve the transfer of phosphate groups from ATP to form phosphorylated intermediates. The second of these intermediates, fructose 1,6-bisphosphate, has a free energy substantially higher than that of glucose. Later in the pathway, these phosphate groups are transferred to ADP to make new molecules of ATP. Although both of these steps use ATP as a substrate, each is catalyzed by a different, specific enzyme.

In reaction 1, the enzyme hexokinase catalyzes the transfer of a phosphate group from ATP to glucose, forming the sugar phosphate glucose 6-phosphate.

In reaction 2, the six-membered glucose ring is rearranged into a five-membered fructose ring.

In reaction 3, the enzyme phosphofructokinase adds a second phosphate to the fructose ring, forming fructose 1,6-bisphosphate.

Reaction 4 opens up the ring and cleaves it to produce two different three-carbon sugar (triose) phosphates: dihydroxyacetone phosphate and glyceraldehyde 3-phosphate.

In reaction 5, one of those products, dihydroxyacetone phosphate, is converted into a second molecule of the other, glyceraldehyde 3-phosphate (G3P).

In summary, by the halfway point of the glycolytic pathway, two things have happened:

- Two molecules of ATP have been invested.
- The six-carbon glucose molecule has been converted into two molecules of a three-carbon sugar phosphate, glyceraldehyde 3-phosphate (G3P).

The energy-harvesting reactions 6–10 of glycolysis yield NADH and ATP

In the discussion that follows, remember that each reaction occurs twice for each glucose molecule because each glucose molecule has been split into two molecules of G3P. The transformation of G3P generates both NADH and ATP. Again, follow the sequence by referring to Figure 9.5.

PRODUCING NADH *Reaction 6* is catalyzed by the enzyme triose phosphate dehydrogenase, and its end product is a phosphate

ester, 1,3-bisphosphoglycerate (BPG). This is an exergonic oxidation reaction, and it is accompanied by a large drop in free energy—more than 100 kcal of energy is released per mole of glucose (**Figure 9.6, left**). The free energy released in this reaction is not lost to heat, but is captured by the accompanying reduction reaction. For each molecule of G3P that is oxidized, one molecule of NAD^+ is reduced to make a molecule of NADH.

NAD^+ is present in only small amounts in the cell, and it must be recycled to allow glycolysis to continue. As we will see, NADH is oxidized back to NAD^+ in the metabolic pathways that follow glycolysis.

PRODUCING ATP *In reactions 7–10* of glycolysis, the two phosphate groups of BPG are transferred one at a time to molecules of ADP, with a rearrangement in between. More than 20 kcal (83.6 kJ/mol) of free energy is stored in ATP for every mole of BPG broken down. Finally, we are left with two moles of pyruvate for every mole of glucose that entered glycolysis.

The enzyme-catalyzed transfer of phosphate groups from donor molecules to ADP to form ATP is called **substrate-level phosphorylation**. (Phosphorylation is the addition of a phosphate group to a molecule.) Substrate-level phosphorylation is distinct from oxidative phosphorylation, which is carried out by the respiratory chain and ATP synthase, and will be discussed later in this chapter. Reaction 7 is an example of substrate-level phosphorylation, in which phosphoglycerate kinase catalyzes the transfer of a phosphate group from BPG to ADP, forming ATP. It is exergonic, even though a substantial amount of energy is consumed in the formation of ATP.

To summarize:

- The energy-investing steps of glycolysis use the energy of hydrolysis of two ATP molecules per glucose molecule.
- The energy-releasing steps of glycolysis produce four ATP molecules per glucose molecule, so the net production of ATP is two molecules.
- The energy-releasing steps of glycolysis produce two molecules of NADH.

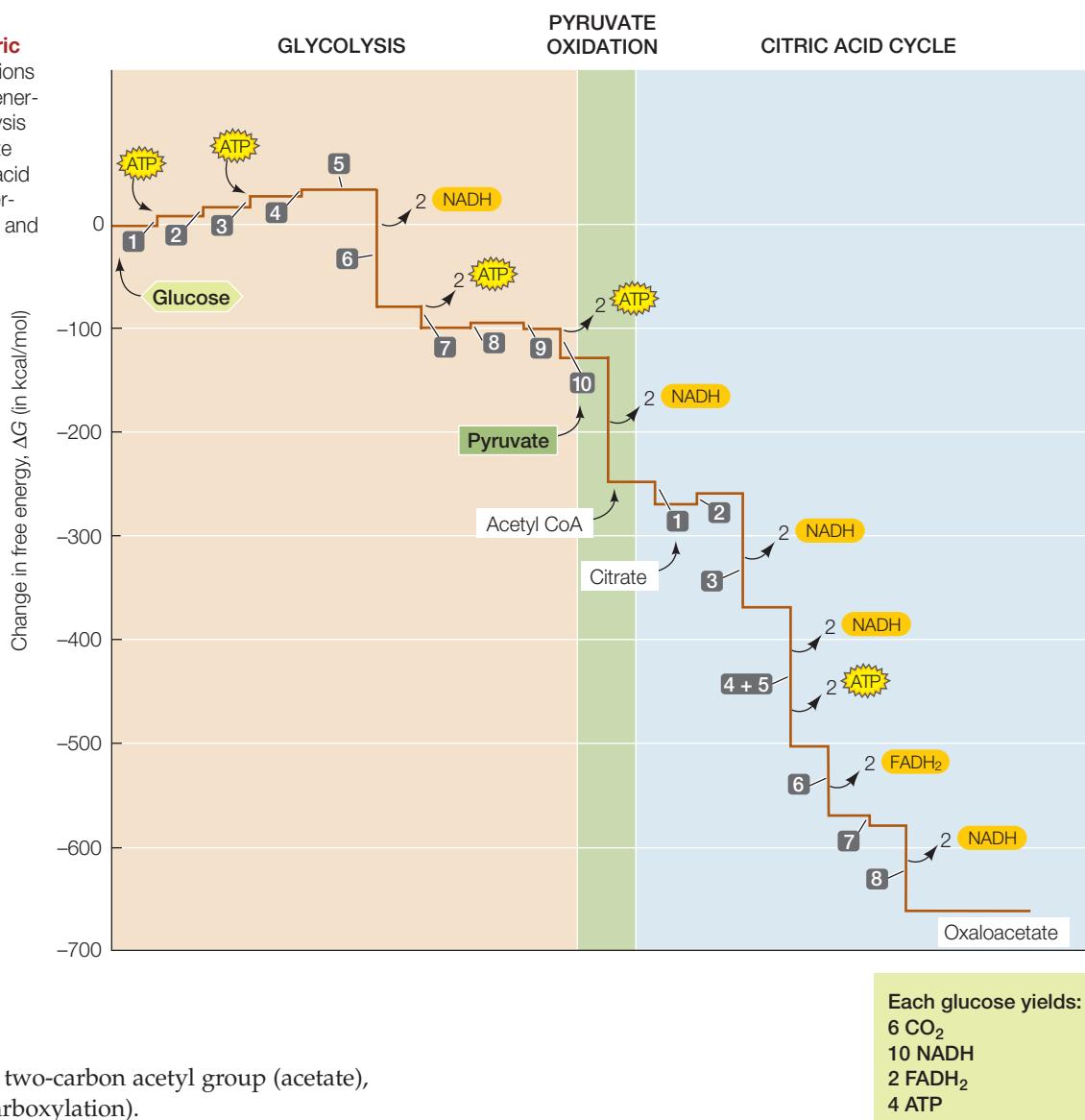
If O_2 is present, glycolysis is followed by the three stages of cellular respiration: pyruvate oxidation, the citric acid cycle, and the respiratory chain/ATP synthesis.

Pyruvate oxidation links glycolysis and the citric acid cycle

In the process of **pyruvate oxidation**, pyruvate is oxidized to the two-carbon acetate molecule, which is then converted to acetyl CoA. This is the link between glycolysis and all the other reactions of cellular respiration. **Coenzyme A (CoA)** is a complex molecule responsible for binding the two-carbon acetate molecule. Acetyl CoA formation is a multi-step reaction catalyzed by the pyruvate dehydrogenase complex, an enormous complex containing 60 individual proteins and 5 different coenzymes. In eukaryotic cells, pyruvate dehydrogenase is located in the mitochondrial matrix (see Figure 5.12). Pyruvate enters the mitochondrion by active transport, and then a series of coupled reactions takes place:

9.6 Changes in Free Energy During Glycolysis and the Citric Acid Cycle

The first five reactions of glycolysis (left) consume free energy, and the remaining five glycolysis reactions release energy. Pyruvate oxidation (middle) and the citric acid cycle (right) both release considerable energy. Refer to Figures 9.5 and 9.7 for the reaction numbers.



- Pyruvate is oxidized to a two-carbon acetyl group (acetate), and CO₂ is released (decarboxylation).
- Part of the energy from this oxidation is captured by the reduction of NAD⁺ to NADH.
- Some of the remaining energy is stored temporarily by combining the acetyl group with CoA, forming acetyl CoA:
 $\text{pyruvate} + \text{NAD}^+ + \text{CoA} + \text{H}^+ \rightarrow \text{acetyl CoA} + \text{NADH} + \text{CO}_2$

(In this reaction, the proton and electrons used to reduce NAD⁺ are derived from the oxidation of both pyruvate and CoA.) Acetyl CoA has 7.5 kcal/mol (31.4 kJ/mol) more energy than simple acetate. Acetyl CoA can donate its acetyl group to various acceptor molecules, much as ATP can donate phosphate groups to various acceptors. But the main role of acetyl CoA is to donate its acetyl group to the four-carbon compound oxaloacetate, forming the six-carbon molecule citrate. This initiates the citric acid cycle, one of life's most important energy-harvesting pathways.

Arsenic, the classic poison of rodent exterminators and murder mysteries, acts by inhibiting pyruvate dehydrogenase, thus decreasing acetyl CoA production. The lack of acetyl CoA stops the citric acid cycle and all the subsequent reactions that de-

pend on it. Consequently, cells eventually run out of ATP and cannot perform essential processes that are powered by ATP hydrolysis.

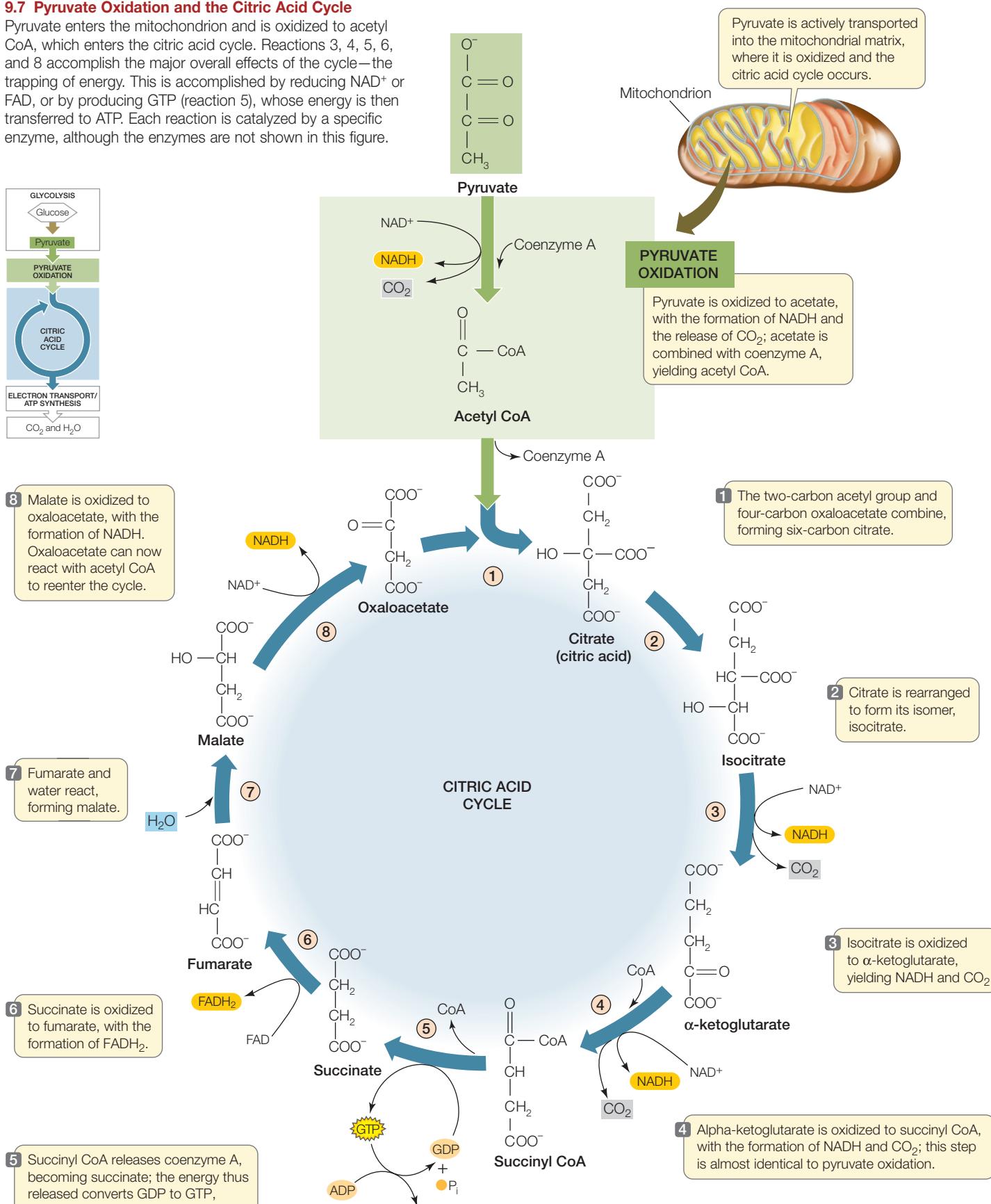
The citric acid cycle completes the oxidation of glucose to CO₂

Acetyl CoA is the starting point for the **citric acid cycle**. This pathway of eight reactions completely oxidizes the two-carbon acetyl group to two molecules of carbon dioxide. The free energy released from these reactions is captured by ADP and the electron carriers NAD⁺ and FAD. **Figure 9.6 right** shows the free energy changes during each step of the pathway.

The citric acid cycle is maintained in a steady state—that is, although the intermediate compounds in the cycle enter and leave it, the concentrations of those intermediates do not change much. Refer to the numbered reactions in **Figure 9.7** as you read the description of each reaction.

9.7 Pyruvate Oxidation and the Citric Acid Cycle

Pyruvate enters the mitochondrion and is oxidized to acetyl CoA, which enters the citric acid cycle. Reactions 3, 4, 5, 6, and 8 accomplish the major overall effects of the cycle—the trapping of energy. This is accomplished by reducing NAD⁺ or FAD, or by producing GTP (reaction 5), whose energy is then transferred to ATP. Each reaction is catalyzed by a specific enzyme, although the enzymes are not shown in this figure.



In reaction 1, the energy temporarily stored in acetyl CoA drives the formation of citrate from oxaloacetate. During this reaction, the CoA molecule is removed and can be reused by pyruvate dehydrogenase.

In reaction 2, the citrate molecule is rearranged to form isocitrate.

In reaction 3, a CO₂ molecule, a proton, and two electrons are removed, converting isocitrate into α-ketoglutarate. This reaction releases a large amount of free energy, some of which is stored in NADH.

In reaction 4, α-ketoglutarate is oxidized to succinyl CoA. This reaction is similar to the oxidation of pyruvate to form acetyl CoA. Like that reaction, it is catalyzed by a multi-enzyme complex and produces CO₂ and NADH.

In reaction 5, some of the energy in succinyl CoA is harvested to make GTP (guanosine triphosphate) from GDP and P_i. This is another example of substrate-level phosphorylation. GTP is then used to make ATP from ADP and P_i.

In reaction 6, the succinate released from succinyl CoA in reaction 5 is oxidized to fumarate. In the process, free energy is released and two hydrogens are transferred to the electron carrier FAD, forming FADH₂.

Reaction 7 is a molecular rearrangement in which water is added to fumarate, forming malate.

In reaction 8, one more NAD⁺ reduction occurs, producing oxaloacetate from malate. Reactions 7 and 8 illustrate a common biochemical mechanism: in reaction 7, water (H₂O) is added to form a hydroxyl (—OH) group, and then in reaction 8 the H from the hydroxyl group is removed, generating a carbonyl group and reducing NAD⁺ to NADH.

The final product, oxaloacetate, is ready to combine with another acetyl group from acetyl CoA and go around the cycle again. The citric acid cycle operates twice for each glucose molecule that enters glycolysis (once for each pyruvate that enters the mitochondrion).

To summarize:

- The *inputs* to the citric acid cycle are acetate (in the form of acetyl CoA), water, and the oxidized electron carriers NAD⁺, FAD, and GDP.
- The *outputs* are carbon dioxide, reduced electron carriers (NADH and FADH₂), and a small amount of GTP. Overall, the citric acid cycle releases two carbons as CO₂ and produces four reduced electron carrier molecules.

The citric acid cycle is regulated by the concentrations of starting materials

We have seen how pyruvate, a three-carbon molecule, is completely oxidized to CO₂ by pyruvate dehydrogenase and the citric acid cycle. For the cycle to continue, the starting molecules—acetyl CoA and oxidized electron carriers—must all be replenished. The electron carriers are reduced during the cycle and in reaction 6 of glycolysis (see Figure 9.5), and they must be reoxidized:



The oxidation of these electron carriers take place in coupled redox reactions, in which other molecules get reduced. When it is present, O₂ is the molecule that eventually accepts these electrons and gets reduced to form H₂O.

9.2 RECAP

The oxidation of glucose in the presence of O₂ involves glycolysis, pyruvate oxidation, and the citric acid cycle. In glycolysis, glucose is converted to pyruvate with some energy capture. Following the initial oxidation of pyruvate, the citric acid cycle completes its oxidation to CO₂ and more energy is captured in the form of reduced electron carriers.

- What is the net energy yield of glycolysis in terms of energy invested and energy harvested? See p. 174 and Figure 9.6
- What role does pyruvate oxidation play in the citric acid cycle? See pp. 174–175 and Figure 9.7
- Explain why reoxidation of NADH is crucial for the continuation of the citric acid cycle. See p. 177

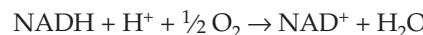
Pyruvate oxidation and the citric acid cycle cannot continue operating unless O₂ is available to receive electrons during the reoxidation of reduced electron carriers. However, these electrons are not passed directly to O₂, as you will learn next.

9.3 How Does Oxidative Phosphorylation Form ATP?

The overall process of ATP synthesis resulting from the reoxidation of electron carriers in the presence of O₂ is called **oxidative phosphorylation**. Two components of the process can be distinguished:

1. *Electron transport*. The electrons from NADH and FADH₂ pass through the **respiratory chain**, a series of membrane-associated electron carriers. The flow of electrons along this pathway results in the active transport of protons out of the mitochondrial matrix and across the inner mitochondrial membrane, creating a proton concentration gradient.
2. *Chemiosmosis*. The protons diffuse back into the mitochondrial matrix through a channel protein, **ATP synthase**, which couples this diffusion to the synthesis of ATP. The inner mitochondrial membrane is otherwise impermeable to protons, so the only way for them to follow their concentration gradient is through the channel.

Before we proceed with the details of these pathways, let's consider an important question: Why should the respiratory chain be such a complex process? Why don't cells use the following single step?



The answer is that this reaction would be untamable. It is extremely exergonic—and would be rather like setting off a stick

of dynamite in the cell. There is no biochemical way to harvest that burst of energy efficiently and put it to physiological use (that is, no single metabolic reaction is so endergonic as to consume a significant fraction of that energy in a single step). To control the release of energy during the oxidation of glucose, cells have evolved a lengthy respiratory chain: a series of reactions, each of which releases a small, manageable amount of energy, one step at a time.

The respiratory chain transfers electrons and releases energy

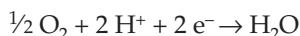
The respiratory chain is located in the inner mitochondrial membrane and contains several interactive components, including large integral proteins, smaller mobile proteins, and a small lipid molecule. **Figure 9.8** shows a plot of the free energy released as electrons are passed between the carriers.

- Four large protein complexes (I, II, III, and IV) contain electron carriers and associated enzymes. In eukaryotes they are integral proteins of the inner mitochondrial membrane (see Figure 5.12), and three are transmembrane proteins.
- Cytochrome *c* is a small peripheral protein that lies in the intermembrane space. It is loosely attached to the outer surface of the inner mitochondrial membrane.
- Ubiquinone (abbreviated Q) is a small, nonpolar, lipid molecule that moves freely within the hydrophobic interior

of the phospholipid bilayer of the inner mitochondrial membrane.

As illustrated in Figure 9.8, NADH passes electrons to protein complex I (called NADH-Q reductase), which in turn passes the electrons to Q. This electron transfer is accompanied by a large drop in free energy. Complex II (succinate dehydrogenase) passes electrons to Q from FADH₂, which was generated in reaction 6 of the citric acid cycle (see Figure 9.7). These electrons enter the chain later than those from NADH and will ultimately produce less ATP.

Complex III (cytochrome *c* reductase) receives electrons from Q and passes them to cytochrome *c*. Complex IV (cytochrome *c* oxidase) receives electrons from cytochrome *c* and passes them to oxygen. Finally the reduction of oxygen to H₂O occurs:



Notice that two protons (H⁺) are also consumed in this reaction. This contributes to the proton gradient across the inner mitochondrial membrane.

During electron transport, protons are also actively transported across the membrane—electron transport within each of the three transmembrane complexes (I, III, and IV) results in the transfer of protons from the matrix to the intermembrane space (**Figure 9.9**). So an imbalance of protons is set up, with the impermeable inner mitochondrial membrane as a barrier. The concentration of H⁺ in the intermembrane space is higher than in the matrix, and this gradient represents a source of potential energy. The diffusion of those protons across the membrane is coupled with the formation of ATP. Thus the energy originally contained in glucose and other fuel molecules is finally captured in the cellular energy currency, ATP. For each pair of electrons passed along the chain from NADH to oxygen, about 2.5 molecules of ATP are formed. FADH₂ oxidation produces about 1.5 ATP molecules.

Electrons from NADH are accepted by NADH-Q reductase at the start of the electron transport chain.

Electrons also come from succinate by way of FADH₂; these electrons are accepted by succinate dehydrogenase.

NADH-Q reductase complex

FADH₂

Succinate dehydrogenase

Ubiquinone (Q)

Cytochrome *c* reductase complex

Cytochrome *c*

Cytochrome *c* oxidase complex

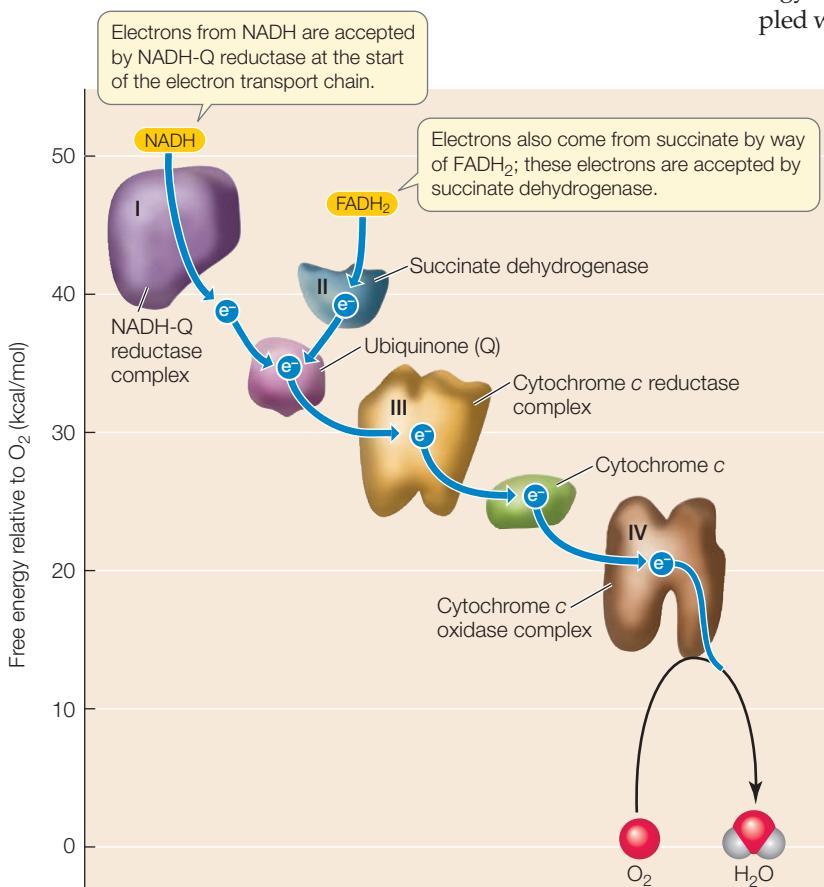
IV

O₂

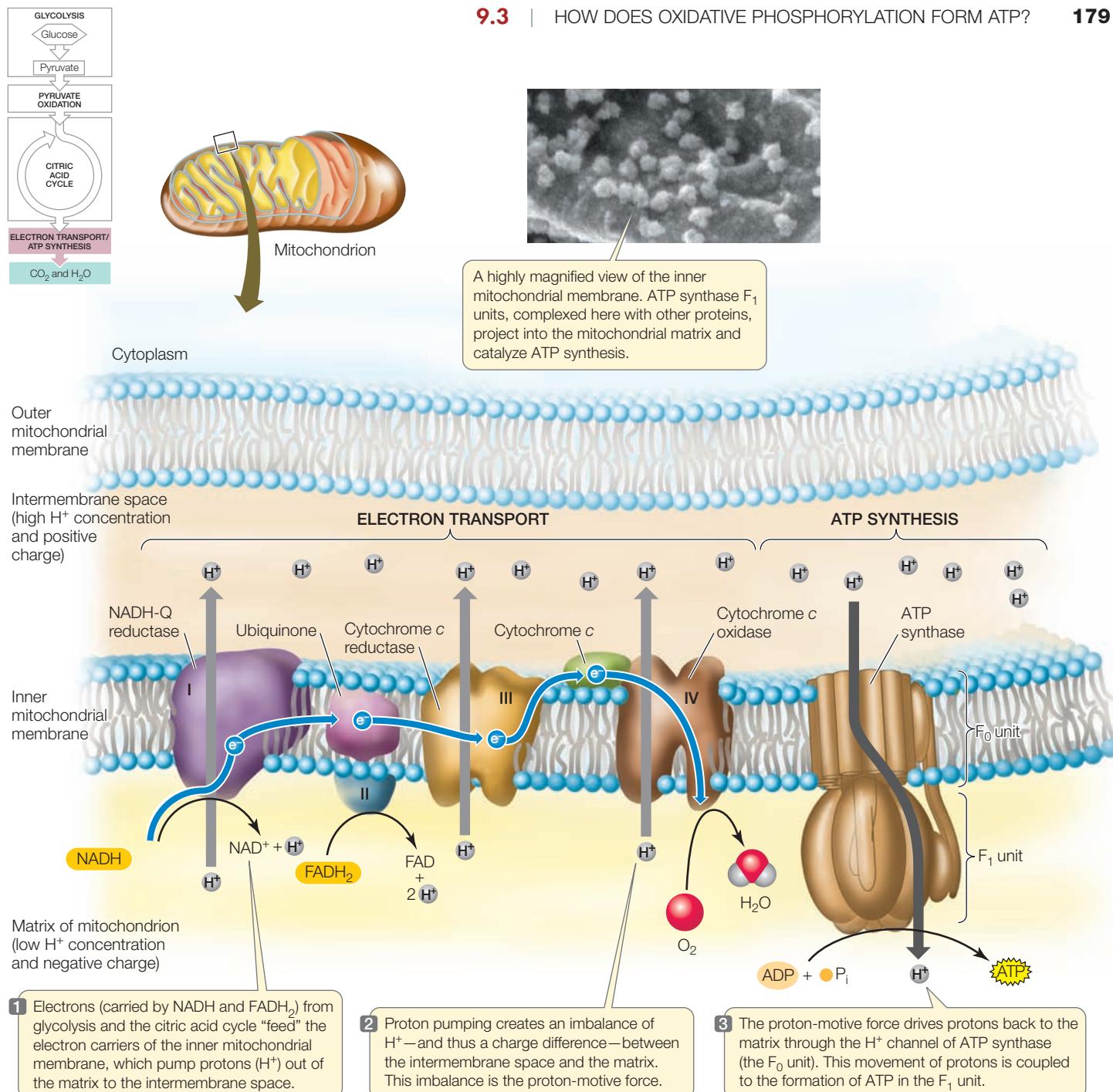
H₂O

Proton diffusion is coupled to ATP synthesis

All the electron carriers and enzymes of the respiratory chain, except cytochrome *c*, are embedded in the inner mitochondrial membrane. As we have just seen, the operation of the respiratory chain results in the active transport of protons from the mitochond-



9.8 The Oxidation of NADH and FADH₂ in the Respiratory Chain Electrons from NADH and FADH₂ are passed along the respiratory chain, a series of protein complexes in the inner mitochondrial membrane containing electron carriers and enzymes. The carriers gain free energy when they become reduced and release free energy when they are oxidized.



9.9 The Respiratory Chain and ATP Synthase Produce ATP by a Chemiosmotic Mechanism

Chemiosmotic Mechanism As electrons pass through the transmembrane protein complexes in the respiratory chain, protons are pumped from the mitochondrial matrix into the intermembrane space. As the protons return to the matrix through ATP synthase, ATP is formed.

yourBioPortal.com
GO TO Animated Tutorial 9.1 • Electron Transport and ATP Synthesis

drial matrix to the intermembrane space. The transmembrane protein complexes (I, III, and IV) act as proton pumps, and as a result, the intermembrane space is more acidic than the matrix.

Because of the positive charge carried by a proton (H⁺), this pumping creates not only a concentration gradient but also a difference in electric charge across the inner mitochondrial

membrane, making the mitochondrial matrix more negative than the intermembrane space. Together, the proton concentration gradient and the electrical charge difference constitute a source of potential energy called the **proton-motive force**. This force tends to drive the protons back across the membrane, just as the charge on a battery drives the flow of electrons to discharge the battery.

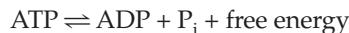
The hydrophobic lipid bilayer is essentially impermeable to protons, so the potential energy of the proton-motive force cannot be discharged by simple diffusion of protons across the membrane. However, protons can diffuse across the membrane by passing through a specific proton channel, called ATP synthase, which couples proton movement to the synthesis of ATP. This coupling of proton-motive force and ATP synthesis is called

the chemiosmotic mechanism (or **chemiosmosis**) and is found in all respiring cells.

THE CHEMIOSMOTIC MECHANISM FOR ATP SYNTHESIS The chemiosmotic mechanism involves transmembrane proteins, including a proton channel and the enzyme ATP synthase, that couple proton diffusion to ATP synthesis. The potential energy of the H⁺

gradient, or the proton-motive force (described above), is harnessed by ATP synthase. This protein complex has two roles: it acts as a channel allowing H⁺ to diffuse back into the matrix, and it uses the energy of that diffusion to make ATP from ADP and P_i.

ATP synthesis is a reversible reaction, and ATP synthase can also act as an ATPase, hydrolyzing ATP to ADP and P_i:



If the reaction goes to the right, free energy is released and is used to pump H⁺ out of the mitochondrial matrix—not the usual mode

yourBioPortal.com

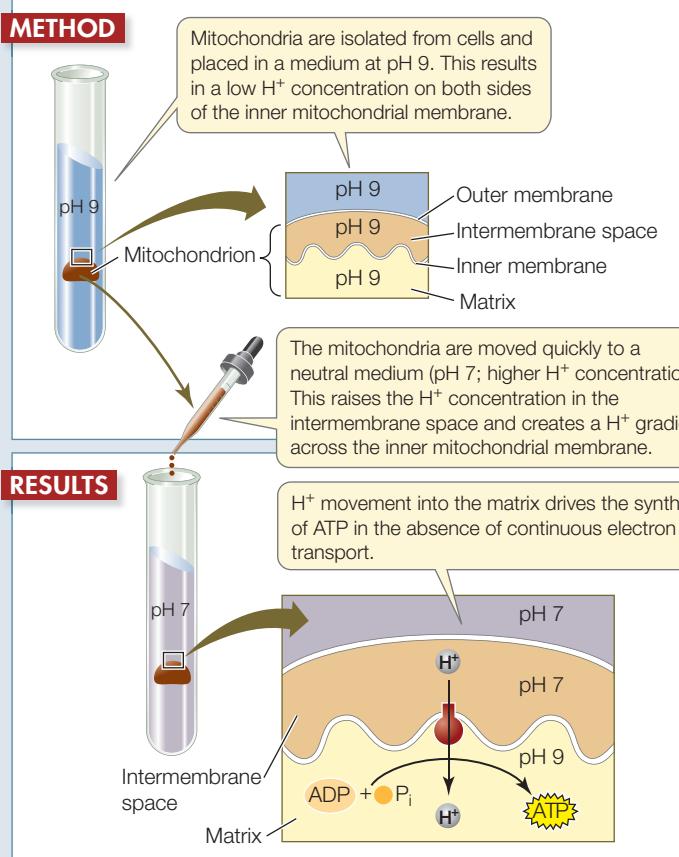
GO TO Animated Tutorial 9.2 • Two Experiments Demonstrate the Chemiosmotic Mechanism

INVESTIGATING LIFE

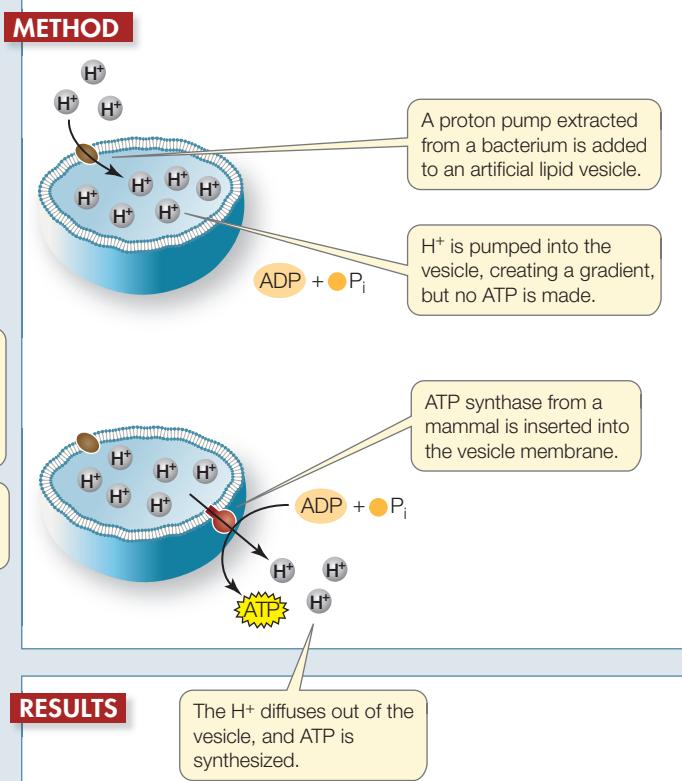
9.10 Two Experiments Demonstrate the Chemiosmotic Mechanism

The chemiosmosis hypothesis was a bold departure for the conventional scientific thinking of the time. It required an intact compartment separated by a membrane. Could a proton gradient drive the synthesis of ATP? And was this capacity entirely due to the ATP synthase enzyme?

HYPOTHESIS A H⁺ gradient can drive ATP synthesis by isolated mitochondria.



HYPOTHESIS ATP synthase is needed for ATP synthesis.



CONCLUSION In the absence of electron transport, an artificial H⁺ gradient is sufficient for ATP synthesis by mitochondria.

CONCLUSION ATP synthase, acting as a H⁺ channel, is necessary for ATP synthesis.

FURTHER INVESTIGATION: What would happen in the experiment on the right if a second ATP synthase, oriented in the opposite way to the one originally inserted in the membrane, were added?

Go to yourBioPortal.com for original citations, discussions, and relevant links for all INVESTIGATING LIFE figures.

of operation. If the reaction goes to the left, it uses the free energy from H^+ diffusion into the matrix to make ATP. What makes it prefer ATP synthesis? There are two answers to this question:

- ATP leaves the mitochondrial matrix for use elsewhere in the cell as soon as it is made, keeping the ATP concentration in the matrix low, and driving the reaction toward the left.
- The H^+ gradient is maintained by electron transport and proton pumping.

Every day a person hydrolyzes about 10^{25} ATP molecules to ADP. This amounts to 9 kg, a significant fraction of almost everyone's entire body weight! The vast majority of this ADP is "recycled"—converted back to ATP—using free energy from the oxidation of glucose.

EXPERIMENTS DEMONSTRATE CHEMIOSMOSIS When it was first proposed almost a half-century ago, the idea that a proton gradient was the energy intermediate linking electron transport to ATP synthesis was a departure from the current conventional thinking. Scientists had been searching for a mitochondrial intermediate that they believed would carry energy in much the same way as the ATP produced by substrate level phosphorylation. The search for this intermediate was not successful, and this led to the idea that chemiosmosis was the mechanism of oxidative phosphorylation. Experimental evidence was needed to support this hypothesis. Two key experiments demonstrated (1) that a proton (H^+) gradient across a membrane can drive ATP synthesis; and (2) that the enzyme ATP synthase is the catalyst for this reaction (Figure 9.10).

In the first experiment, mitochondria without a food source were "fooled" into making ATP by raising the H^+ concentration in their environment. In the second experiment, a light-driven proton pump isolated from bacteria was inserted into artificial lipid vesicles. This generated a proton gradient, but since ATP synthase was absent, ATP was not made. Then, ATP synthase was inserted into the vesicles and ATP was generated.

UNCOUPLING PROTON DIFFUSION FROM ATP PRODUCTION The tight coupling between H^+ diffusion and the formation of ATP provides further evidence for the chemiosmotic mechanism. If a second type of H^+ diffusion channel (that does not synthesize ATP) is present in the mitochondrial membrane, the energy of the H^+ gradient is released as heat rather than being coupled to ATP synthesis. Such uncoupling molecules actually exist in the mitochondria of some organisms to generate heat instead of ATP. For example, the natural uncoupling protein thermogenin plays an important role in regulating the temperatures of newborn human infants, who lack hair to keep warm, and in hibernating animals.

A popular weight loss drug in the 1930s was the uncoupler molecule, dinitrophenol. There were claims of dramatic weight loss when the drug was administered to obese patients. Unfortunately, the heat that was released caused fatally high fevers, and the effective dose and fatal dose were quite close. So the use of this drug was discontinued in 1938. However, the general strategy of using an uncoupler for weight loss remains a subject of research.

HOW ATP SYNTHASE WORKS: A MOLECULAR MOTOR Now that we have established that the H^+ gradient is needed for ATP synthesis, a question remains: how does the enzyme actually make ATP from ADP and P_i ? This is certainly a fundamental question in biology, as it underlies energy harvesting in most cells. Look at the structure of ATP synthase in Figure 9.9. It is a molecular motor composed of two parts: the F_0 unit, a transmembrane region that is the H^+ channel, and the F_1 unit, which contains the active sites for ATP synthesis. F_1 consists of six subunits (three each of two polypeptide chains), arranged like the segments of an orange around a central polypeptide. ATP synthesis is coupled with conformational changes in the ATP synthase enzyme, which are induced by proton movement through the complex. The potential energy set up by the proton gradient across the inner membrane drives the passage of protons through the ring of polypeptides that make up the F_0 component. This ring rotates as the protons pass through the membrane, causing the F_1 unit to rotate as well. ADP and P_i bind to active sites that become exposed on the F_1 unit as it rotates, and ATP is made. The structure and function of ATP synthase are shared by living organisms as diverse as bacteria and humans. These molecular motors make ATP at rates up to 100 molecules per second.

9.3 RECAP

The oxidation of reduced electron carriers in the respiratory chain drives the active transport of protons across the inner mitochondrial membrane, generating a proton-motive force. Diffusion of protons down their electrochemical gradient through ATP synthase is coupled to the synthesis of ATP.

- What are the roles of oxidation and reduction in the respiratory chain? See Figures 9.8 and 9.9
- What is the proton motive force and how does it drive chemiosmosis? See pp. 179–180
- Explain how the two experiments described in Figure 9.10 demonstrate the chemiosmotic mechanism. See p. 181

Oxidative phosphorylation captures a great deal of energy in ATP. But it does not occur if O_2 is absent. We turn now to the metabolism of glucose in anaerobic conditions.

9.4 How Is Energy Harvested from Glucose in the Absence of Oxygen?

In the absence of O_2 (anaerobic conditions), a small amount of ATP can be produced by glycolysis and fermentation. Like glycolysis, fermentation pathways occur in the cytoplasm. There are many different types of fermentation, but they all operate to regenerate NAD^+ so that the NAD-requiring reaction of glycolysis can continue (see reaction 6 in Figure 9.5). Of course, if a necessary reactant such as NAD^+ is not present, the reaction will not take place. How do fermentation reactions regenerate NAD^+ and permit ATP formation to continue?

Prokaryotic organisms often live in O₂-deficient environments and are known to use many different fermentation pathways. But the two best understood fermentation pathways are found in a wide variety of organisms including eukaryotes. These two short pathways are lactic acid fermentation, whose end product is lactic acid (lactate); and alcoholic fermentation, whose end product is ethyl alcohol (ethanol).

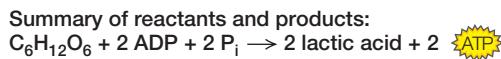
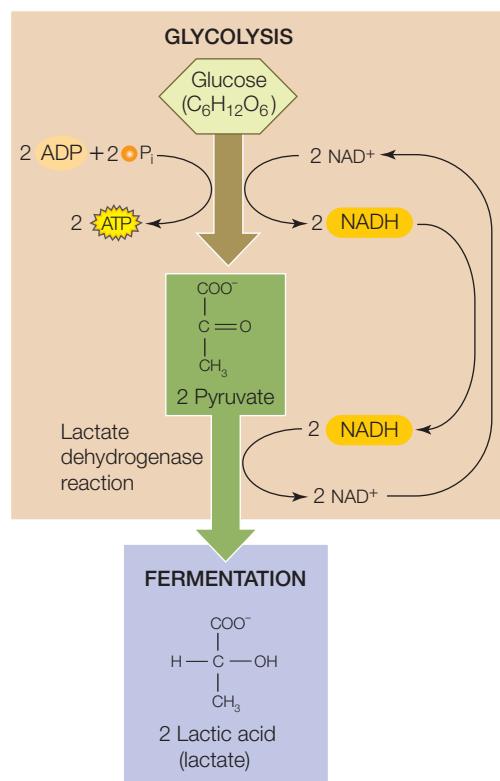
In lactic acid fermentation, pyruvate serves as the electron acceptor and lactate is the product (Figure 9.11). This process takes place in many microorganisms and complex organisms, including higher plants and vertebrates. A notable example of lactic acid fermentation occurs in vertebrate muscle tissue. Usually, vertebrates get their energy for muscle contraction aerobically, with the circulatory system supplying O₂ to muscles. In small vertebrates, this is almost always adequate: for example, birds can fly long distances without resting. But in larger vertebrates such as humans, the circulatory system is not up to the task of delivering enough O₂ when the need is great, such as during high activity. At this point, the muscle cells break down glycogen (a stored polysaccharide) and undergo lactic acid fermentation.

Lactic acid buildup becomes a problem after prolonged periods because the acid ionizes, forming H⁺ and lowering the pH of the cell. This affects cellular activities and causes muscle

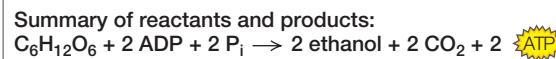
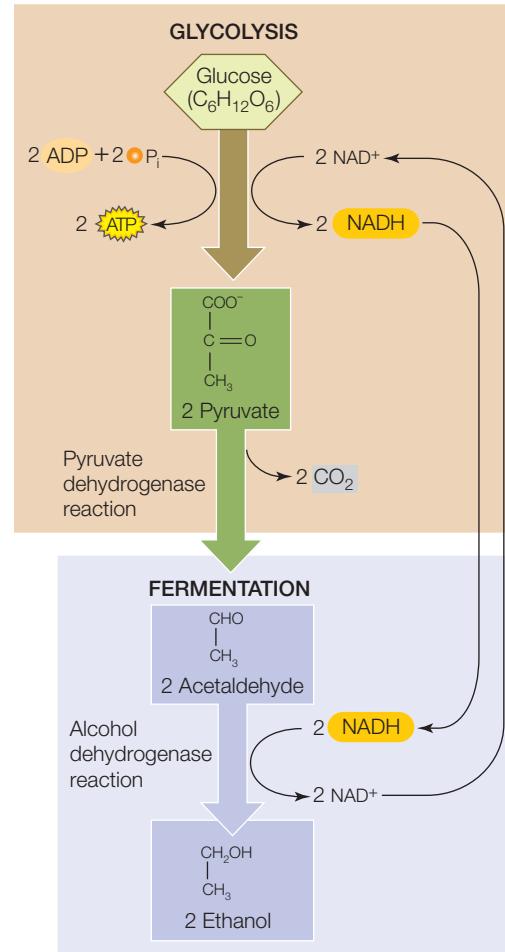
cramps, resulting in muscle pain, which abates upon resting. Lactate dehydrogenase, the enzyme that catalyzes the fermentation reaction, works in both directions. That is, it can catalyze the oxidation of lactate as well as the reduction of pyruvate. When lactate levels are decreased, muscle activity can resume.

Alcoholic fermentation takes place in certain yeasts (eukaryotic microbes) and some plant cells under anaerobic conditions. This process requires two enzymes, pyruvate dehydrogenase and alcohol dehydrogenase, which metabolize pyruvate to ethanol (Figure 9.12). As with lactic acid fermentation, the reactions are essentially reversible. For thousands of years, humans have used anaerobic fermentation by yeast cells to produce alcoholic beverages. The cells use sugars from plant sources (glucose from grapes or maltose from barley) to produce the end product, ethanol, in wine and beer.

By recycling NAD⁺, fermentation allows glycolysis to continue, thus producing small amounts of ATP through substrate-level phosphorylation. The net yield of two ATPs per glucose



9.11 Lactic Acid Fermentation Glycolysis produces pyruvate, ATP, and NADH from glucose. Lactic acid fermentation uses NADH as a reducing agent to reduce pyruvate to lactic acid (lactate), thus regenerating NAD⁺ to keep glycolysis operating.



9.12 Alcoholic Fermentation In alcoholic fermentation, pyruvate from glycolysis is converted into acetaldehyde, and CO₂ is released. NADH from glycolysis is used to reduce acetaldehyde to ethanol, thus regenerating NAD⁺ to keep glycolysis operating.

molecule is much lower than the energy yield from cellular respiration. For this reason, most organisms existing in anaerobic environments are small microbes that grow relatively slowly.

Cellular respiration yields much more energy than fermentation

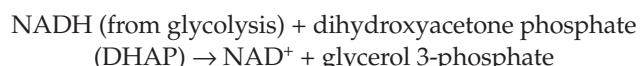
The total net energy yield from glycolysis plus fermentation is two molecules of ATP per molecule of glucose oxidized. The maximum yield of ATP that can be harvested from a molecule of glucose through glycolysis followed by cellular respiration is much greater—about 32 molecules of ATP (Figure 9.13). (Review Figures 9.5, 9.7, and 9.9 to see where all the ATP molecules come from.)

Why do the metabolic pathways that operate in aerobic environments produce so much more ATP? Glycolysis and fermentation only partially oxidize glucose, as does fermentation. Much more energy remains in the end products of fermentation (lactic acid and ethanol) than in CO_2 , the end product of cellular respiration. In cellular respiration, carriers (mostly NAD^+) are reduced in pyruvate oxidation and the citric acid cycle. Then the reduced carriers are oxidized by the respiratory chain, with the accompanying production of ATP by chemiosmosis (2.5 ATP for each NADH and 1.5 ATP for each FADH_2). In an aerobic environment, a cell or organism capable of aerobic metabolism will have the advantage over one that is limited to fermentation, in terms of its ability to harvest chemical energy. Two key events in the evolution of multicellular organisms were the rise in atmospheric O_2 levels (see Chapter 1) and the development of metabolic pathways to utilize that O_2 .

The yield of ATP is reduced by the impermeability of some mitochondria to NADH

The total gross yield of ATP from the oxidation of one molecule of glucose to CO_2 is 32. However, in some animal cells the inner mitochondrial membrane is impermeable to NADH, and a “toll” of one ATP must be paid for each NADH molecule that is produced in glycolysis and must be “shuttled” into the mitochondrial matrix. So in these animals, the net yield of ATP is 30.

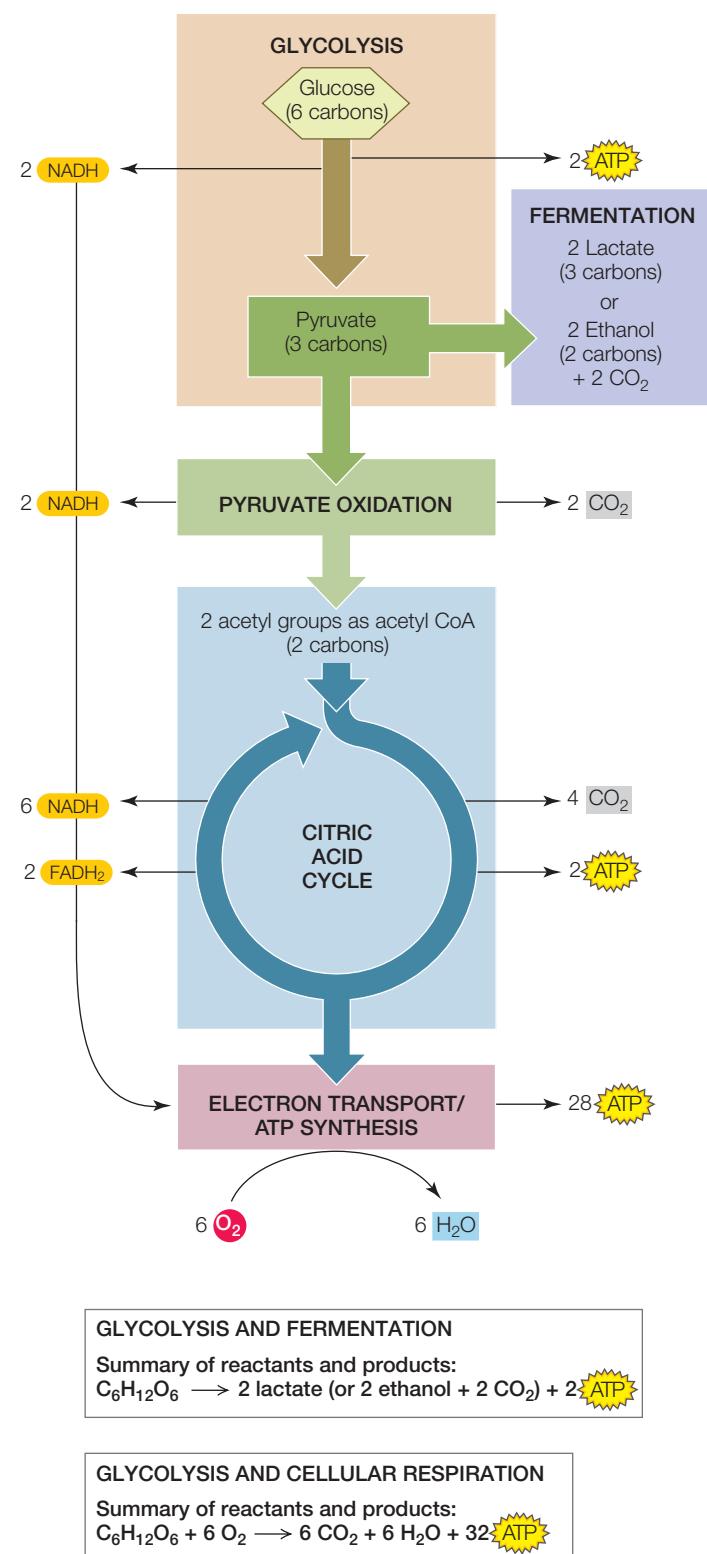
NADH shuttle systems transfer the electrons captured by glycolysis onto substrates that are capable of movement across the mitochondrial membranes. In muscle and liver tissues, an important shuttle involves glycerol 3-phosphate. In the cytosol,



Glycerol 3-phosphate crosses both mitochondrial membranes. In the mitochondrial matrix,



DHAP is able to move back to the cytosol, where it is available to repeat the process. Note that the reducing electrons are transferred from NADH outside the mitochondrion to FADH_2 inside the mitochondrion. As you know from Figures 9.8 and 9.9, the energy yield in terms of ATP from FADH_2 is lower than that from NADH. This lowers the overall energy yield.



9.13 Cellular Respiration Yields More Energy Than Fermentation

Electron carriers are reduced in pyruvate oxidation and the citric acid cycle, then oxidized by the respiratory chain. These reactions produce ATP via chemiosmosis.

9.4 RECAP

In the absence of O_2 , fermentation pathways use NADH formed by glycolysis to reduce pyruvate and regenerate NAD⁺. The energy yield of fermentation is low because glucose is only partially oxidized. When O_2 is present, the electron carriers of cellular respiration allow for the full oxidation of glucose, so the energy yield from glucose is much higher.

- Why is replenishing NAD⁺ crucial to cellular metabolism?
See pp. 182–183
- What is the total energy yield from glucose in human cells in the presence versus the absence of O_2 ?
See p. 183 and Figure 9.13

Now that you've seen how cells harvest energy, let's see how that energy moves through other metabolic pathways in the cell.

9.5 How Are Metabolic Pathways Interrelated and Regulated?

Glycolysis and the pathways of cellular respiration do not operate in isolation. Rather, there is an interchange of molecules into and out of these pathways, to and from the metabolic pathways for the synthesis and breakdown of amino acids, nucleotides, fatty acids, and other building blocks of life. Carbon skeletons can enter the catabolic pathways and be broken down to release their energy, or they can enter anabolic pathways to be used in the formation of the macromolecules that are the major constituents of the cell. These relationships are summarized in **Figure 9.14**. In this section we will explore how pathways are interrelated by the sharing of intermediate substances, and we will see how pathways are regulated by the inhibitors of key enzymes.

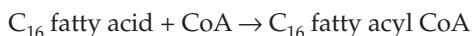
Catabolism and anabolism are linked

A hamburger or veggie burger on a bun contains three major sources of carbon skeletons: carbohydrates, mostly in the form of starch (a polysaccharide); lipids, mostly as triglycerides (three fatty acids attached to glycerol); and proteins (polymers of amino acids). Look at Figure 9.14 to see how each of these three types of macromolecules can be hydrolyzed and used in catabolism or anabolism.

9.14 Relationships among the Major Metabolic Pathways of the Cell Note the central positions of glycolysis and the citric acid cycle in this network of metabolic pathways. Also note that many of the pathways can operate essentially in reverse.

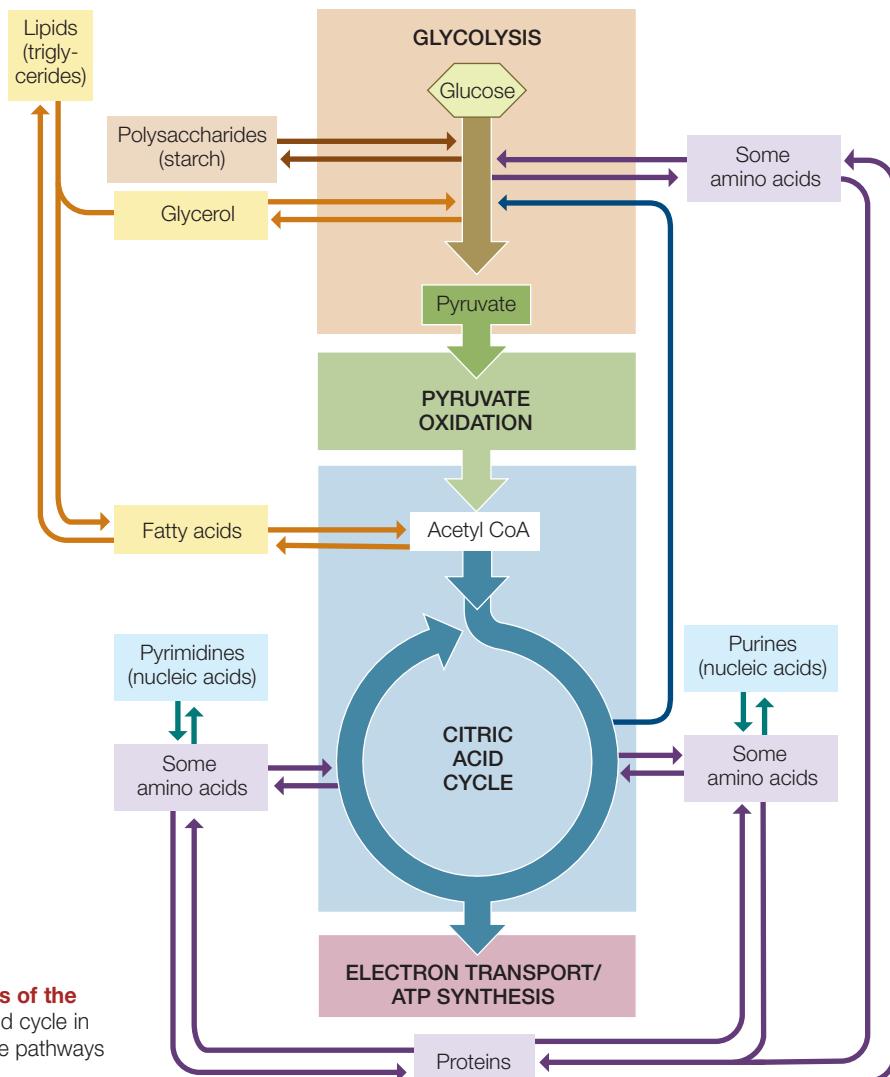
CATABOLIC INTERCONVERSIONS Polysaccharides, lipids, and proteins can all be broken down to provide energy:

- *Polysaccharides* are hydrolyzed to glucose. Glucose then passes through glycolysis and cellular respiration, where its energy is captured in ATP.
- *Lipids* are broken down into their constituents, glycerol and fatty acids. Glycerol is converted into dihydroxyacetone phosphate (DHAP), an intermediate in glycolysis. Fatty acids are highly reduced molecules that are converted to acetyl CoA inside the mitochondrion by a series of oxidation enzymes, in a process known as β -oxidation. For example, the β -oxidation of a C₁₆ fatty acid occurs in several steps:

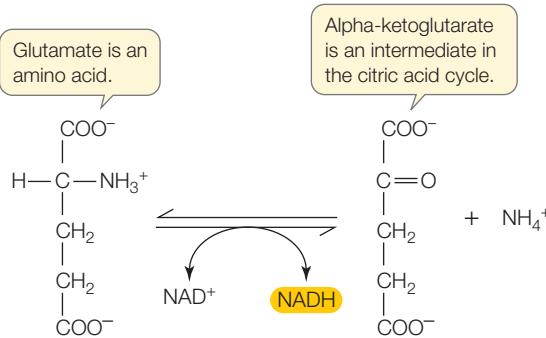


repeat 6 times \rightarrow 8 acetyl CoA

The acetyl CoA can then enter the citric acid cycle and be catabolized to CO₂.



- Proteins are hydrolyzed to their amino acid building blocks. The 20 different amino acids feed into glycolysis or the citric acid cycle at different points. For example, the amino acid glutamate is converted into α -ketoglutarate, an intermediate in the citric acid cycle.



ANABOLIC INTERCONVERSIONS Many catabolic pathways can operate essentially in reverse, with some modifications. Glycolytic and citric acid cycle intermediates, instead of being oxidized to form CO_2 , can be reduced and used to form glucose in a process called **gluconeogenesis** (which means “new formation of glucose”). Likewise, acetyl CoA can be used to form fatty acids. The most common fatty acids have even numbers of carbons: 14, 16, or 18. These are formed by the addition of two-carbon acetyl CoA “units” one at a time until the appropriate chain length is reached. Acetyl CoA is also a building block for various pigments, plant growth substances, rubber, steroid hormones, and other molecules.

Some intermediates in the citric acid cycle are reactants in pathways that synthesize important components of nucleic acids. For example, α -ketoglutarate is a starting point for purines, and oxaloacetate for pyrimidines. In addition, α -ketoglutarate is a starting point for the synthesis of chlorophyll (used in photosynthesis; see Chapter 10) and the amino acid glutamate (used in protein synthesis).

Catabolism and anabolism are integrated

A carbon atom from a protein in your burger can end up in DNA, fat, or CO_2 , among other fates. How does the organism “decide” which metabolic pathways to follow, in which cells? With all of the possible interconversions, you might expect that cellular concentrations of various biochemical molecules would vary widely. Remarkably, the levels of these substances in what is called the metabolic pool—the sum total of all the biochemical molecules in a cell—are quite constant. Organisms regulate the enzymes of catabolism and anabolism in various cells in order to maintain a balance. This metabolic homeostasis gets upset only in unusual circumstances. Let’s look one such unusual circumstance: undernutrition.

Glucose is an excellent source of energy, but lipids and proteins can also be broken down and their constituents used as energy sources. Any one or all three of these types of molecules could be used to provide the energy your body needs. But normally these substances are not equally available for energy me-

tabolism and ATP formation. Proteins, for example, have essential roles as enzymes and as structural elements, providing support and movement; they are not stored for energy, and using them for energy might deprive the body of other vital functions.

Fats (triglycerides) do not have catalytic roles. Because they are nonpolar, fats do not bind water, and they are therefore less dense than polysaccharides in aqueous environments. In addition, fats are more reduced than carbohydrates (have more C—H bonds and fewer C—OH bonds) and thus have more energy stored in their bonds (see Figure 9.2). So it is not surprising that fats are the preferred energy store in many organisms. The human body stores fats and carbohydrates; fats are stored in adipose tissue, and glucose is stored as the polysaccharide glycogen in muscles and the liver. A typical person has about one day’s worth of food energy stored as glycogen (a polysaccharide) and over a month’s food energy stored as fats.

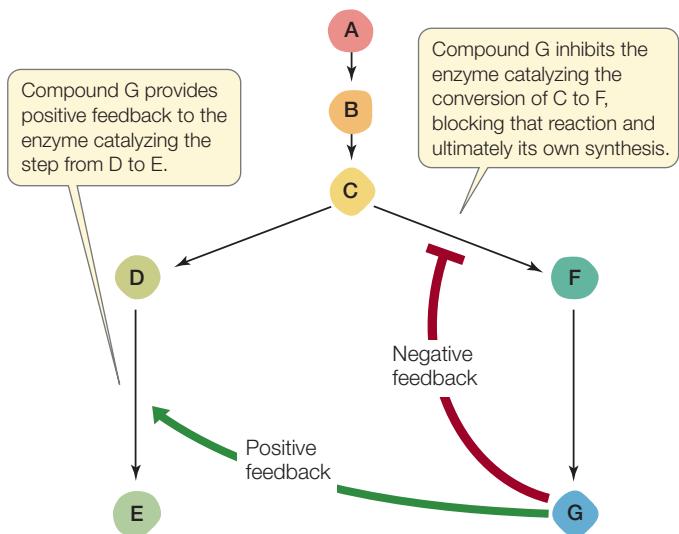
What happens if a person does not eat enough to produce sufficient ATP and NADH for anabolism and biological activities? This situation can be deliberate (to lose weight), but for too many people, it is forced upon them because not enough food is available, resulting in undernutrition and starvation. Initially, homeostasis can be maintained. The first energy stores to be used are the glycogen stores in muscle and liver cells. These stores do not last long, and next come the fats.

In cells that have access to fatty acids, their breakdown produces acetyl CoA for cellular respiration. However, a problem remains: because fatty acids cannot cross from the blood to the brain, the brain can use only glucose as its energy source. With glycogen already depleted, the body must convert something else to make glucose for the brain. This is accomplished by the breakdown of proteins and the conversion of their amino acids to glucose by gluconeogenesis. Without sufficient food intake, proteins and fats are used up. After several weeks of starvation, fat stores become depleted, and the only energy source left is protein. At this point, essential structural proteins, enzymes, and antibodies get broken down. The loss of such proteins can lead to severe illness and eventual death.

Metabolic pathways are regulated systems

We have described the relationships between metabolic pathways and noted that these pathways work together to provide homeostasis in the cell and organism. But how does the cell regulate the interconversions between pathways to maintain constant metabolic pools? This is a problem of systems biology, which seeks to understand how biochemical pathways interact (see Figure 8.15). It is a bit like trying to predict traffic patterns in a city: if an accident blocks traffic on a major road, drivers take alternate routes, where the traffic volume consequently changes.

Consider what happens to the starch in your burger bun. In the digestive system, starch is hydrolyzed to glucose, which enters the blood for distribution to the rest of the body. But before the glucose is distributed, a regulatory check must be made: if there is already enough glucose in the blood to supply the body’s needs, the excess glucose is converted into glycogen and



9.15 Regulation by Negative and Positive Feedback

Allosteric feedback regulation plays an important role in metabolic pathways. The accumulation of some products can shut down their synthesis, or can stimulate other pathways that require the same raw materials.

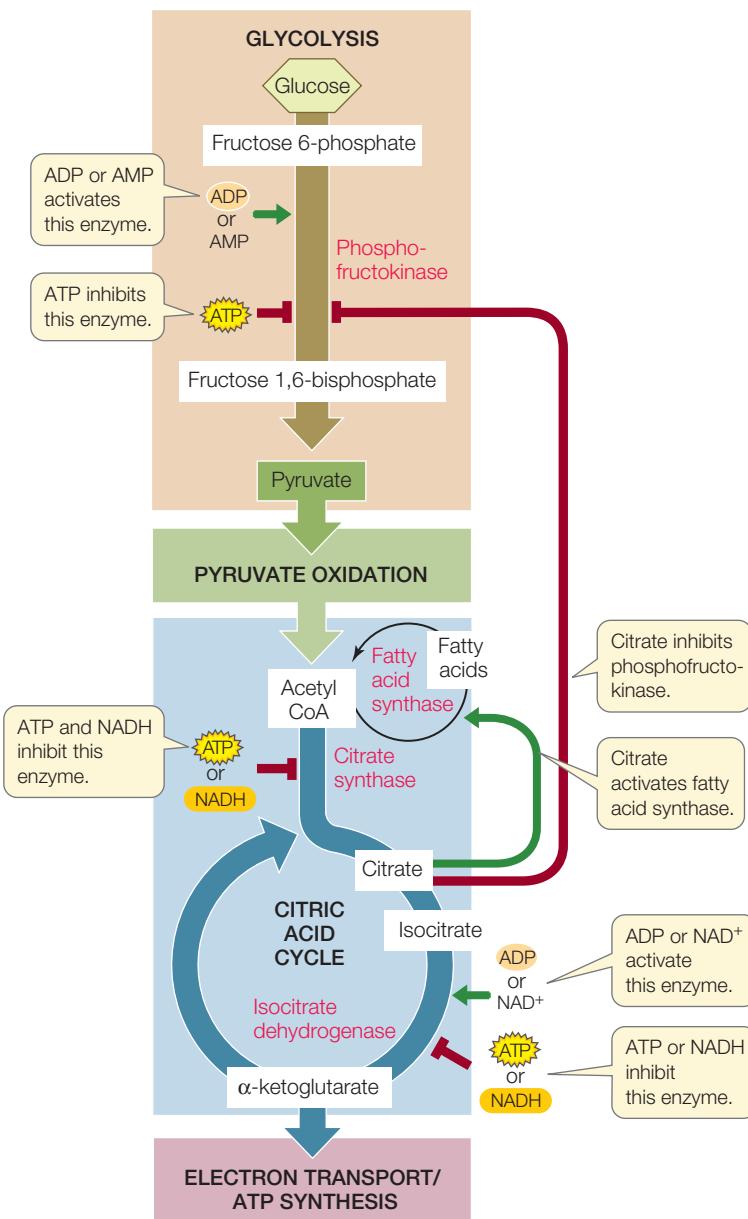
stored in the liver. If not enough glucose is supplied by food, glycogen is broken down, or other molecules are used to make glucose by gluconeogenesis.

The end result is that the level of glucose in the blood is remarkably constant. How does the body accomplish this?

Glycolysis, the citric acid cycle, and the respiratory chain are subject to *allosteric regulation* (see Section 8.5) of the enzymes involved. An example of allosteric regulation is feedback inhibition, illustrated in Figure 8.19. In a metabolic pathway, a high concentration of the final product can inhibit the action of an enzyme that catalyzes an earlier reaction. On the other hand, an excess of the product of one pathway can speed up reactions in another pathway, diverting raw materials away from synthesis of the first product (Figure 9.15). These negative and positive feedback mechanisms are used at many points in the energy-harvesting pathways, and are summarized in Figure 9.16.

- The main control point in glycolysis is the enzyme *phosphofructokinase* (reaction 3 in Figure 9.5). This enzyme is allosterically inhibited by ATP or citrate, and activated by ADP or AMP. Under anaerobic conditions, fermentation yields a relatively small amount of ATP, and phosphofructokinase operates at a high rate. However when conditions are aerobic, respiration makes 16 times more ATP than fermentation does, and the abundant ATP allosterically inhibits phosphofructokinase. Consequently, the conversion of fructose 6-phosphate to fructose 1,6-bisphosphate declines, and so does the rate of glucose utilization.
- The main control point in the citric acid cycle is the enzyme *isocitrate dehydrogenase*, which converts isocitrate to α -ketoglutarate (reaction 3 in Figure 9.7). NADH and

ATP are feedback inhibitors of this reaction, while ADP and NAD⁺ are activators. If too much ATP or NADH accumulates, the conversion of isocitrate is slowed, and the citric acid cycle shuts down. A shutdown of the citric acid cycle would cause large amounts of isocitrate and citrate to accumulate if the production of citrate were not also slowed. But, as mentioned above, an excess of citrate acts as a feedback inhibitor of phosphofructokinase. Thus, if the citric acid cycle has been slowed or shut down because of abun-



9.16 Allosteric Regulation of Glycolysis and the Citric Acid Cycle

Allosteric regulation controls glycolysis and the citric acid cycle at crucial early steps, increasing their efficiency and preventing the excessive buildup of intermediates.

dant ATP (and not because of a lack of oxygen), glycolysis is slowed as well. Both processes resume when the ATP level falls and they are needed again. Allosteric regulation keeps these processes in balance.

- Another control point involves *acetyl CoA*. If the level of ATP is high and the citric acid cycle shuts down, the accumulation of citrate activates fatty acid synthase, diverting acetyl CoA to the synthesis of fatty acids for storage. That is one reason why people who eat too much accumulate fat. These fatty acids may be metabolized later to produce more acetyl CoA.

9.5 RECAP

Glucose can be made from intermediates in glycolysis and the citric acid cycle by a process called gluconeogenesis. The metabolic pathways for the production and breakdown of lipids and amino acids are tied to those of glucose metabolism. Reaction products regulate key enzymes in the various pathways.

- Give examples of a catabolic interconversion of a lipid and of an anabolic interconversion of a protein. [See pp. 184–185 and Figure 9.14](#)
- How does phosphofructokinase serve as a control point for glycolysis? [See p. 186 and Figure 9.16](#)
- Describe what would happen if there was no allosteric mechanism for modulating the level of acetyl CoA.

CHAPTER SUMMARY

9.1 How Does Glucose Oxidation Release Chemical Energy?

- As a material is **oxidized**, the electrons it loses are transferred to another material, which is thereby **reduced**. Such **redox reactions** transfer large amounts of energy. [Review Figure 9.2, WEB ACTIVITIES 9.1 and 9.2](#)
- The coenzyme **NAD⁺** is a key electron carrier in biological redox reactions. It exists in two forms, one oxidized (**NAD⁺**) and the other reduced (**NADH**).
- Glycolysis** operates in the presence or absence of **O₂**. Under **aerobic** conditions, **cellular respiration** continues the process of breaking down glucose. Under **anaerobic** conditions, **fermentation** occurs. [Review Figure 9.4](#)
- The pathways of cellular respiration after glycolysis are **pyruvate oxidation**, the **citric acid cycle**, and the **electron transport/ATP synthesis**.

9.2 What Are the Aerobic Pathways of Glucose Metabolism?

- Glycolysis consists of 10 enzyme-catalyzed reactions that occur in the cell cytoplasm. Two **pyruvate** molecules are produced for each partially oxidized molecule of glucose, providing the starting material for both cellular respiration and fermentation. [Review Figure 9.5](#)
- The first five reactions of glycolysis require an investment of energy; the last five produce energy. The net gain is two molecules of ATP. [Review Figure 9.6](#)
- The enzyme-catalyzed transfer of phosphate groups to ADP by enzymes other than ATPase is called **substrate-level phosphorylation** and produces ATP.
- Pyruvate oxidation follows glycolysis and links glycolysis to the citric acid cycle. This pathway converts pyruvate into **acetyl CoA**.
- Acetyl CoA is the starting point of the citric acid cycle. It reacts with oxaloacetate to produce citrate. A series of eight enzyme-catalyzed reactions oxidize citrate and regenerate oxaloacetate, continuing the cycle. [Review Figure 9.7, WEB ACTIVITY 9.3](#)

9.3 How Does Oxidative Phosphorylation Form ATP?

- Oxidation of electron carriers in the presence of **O₂** releases energy that can be used to form ATP in a process called **oxidative phosphorylation**.
- The NADH and FADH₂ produced in glycolysis, pyruvate oxidation, and the citric acid cycle are oxidized by the respiratory

chain, regenerating NAD⁺ and FAD. Oxygen (O₂) is the final acceptor of electrons and protons, forming water (H₂O). [Review Figure 9.8, WEB ACTIVITY 9.4](#)

- The respiratory chain not only transports electrons, but also pumps protons across the inner mitochondrial membrane, creating the **proton-motive force**.
- Protons driven by the proton-motive force can return to the mitochondrial matrix via **ATP synthase**, a molecular motor that couples this movement of protons to the synthesis of ATP. This process is called **chemiosmosis**. [Review Figure 9.9, ANIMATED TUTORIALS 9.1 and 9.2](#)

9.4 How Is Energy Harvested from Glucose in the Absence of Oxygen?

- In the absence of O₂, glycolysis is followed by fermentation. Together, these pathways partially oxidize pyruvate and generate end products such as **lactic acid** or **ethanol**. In the process, NAD⁺ is regenerated from NADH so that glycolysis can continue, thus generating a small amount of ATP. [Review Figures 9.11 and 9.12](#)
- For each molecule of glucose used, fermentation yields 2 molecules of ATP. In contrast, glycolysis operating with pyruvate oxidation, the citric acid cycle, and the respiratory chain/ATP synthase yields up to 32 molecules of ATP per molecule of glucose. [Review Figure 9.13, WEB ACTIVITY 9.5](#)

9.5 How Are Metabolic Pathways Interrelated and Regulated?

- The **catabolic pathways** for the breakdown of carbohydrates, fats, and proteins feed into the energy-harvesting metabolic pathways. [Review Figure 9.14](#)
- Anabolic pathways** use intermediate components of the energy-harvesting pathways to synthesize fats, amino acids, and other essential building blocks.
- The formation of glucose from intermediates of glycolysis and the citric acid cycle is called **gluconeogenesis**.
- The rates of glycolysis and the citric acid cycle are controlled by **allosteric regulation** and by the diversion of excess acetyl CoA into fatty acid synthesis. Key regulated enzymes include phosphofructokinase, citrate synthase, isocitrate dehydrogenase, and fatty acid synthase. [See Figure 9.16, WEB ACTIVITY 9.6](#)

SELF-QUIZ

1. The role of oxygen gas in our cells is to
 - a. catalyze reactions in glycolysis.
 - b. produce CO_2 .
 - c. form ATP.
 - d. accept electrons from the respiratory chain.
 - e. react with glucose to split water.
2. Oxidation and reduction
 - a. entail the gain or loss of proteins.
 - b. are defined as the loss of electrons.
 - c. are both endergonic reactions.
 - d. always occur together.
 - e. proceed only under aerobic conditions.
3. NAD⁺ is
 - a. a type of organelle.
 - b. a protein.
 - c. present only in mitochondria.
 - d. a part of ATP.
 - e. formed in the reaction that produces ethanol.
4. Glycolysis
 - a. takes place in the mitochondrion.
 - b. produces no ATP.
 - c. has no connection with the respiratory chain.
 - d. is the same thing as fermentation.
 - e. reduces two molecules of NAD⁺ for every glucose molecule processed.
5. Fermentation
 - a. takes place in the mitochondrion.
 - b. takes place in all animal cells.
 - c. does not require O_2 .
 - d. requires lactic acid.
 - e. prevents glycolysis.
6. Which statement about pyruvate is *not* true?
 - a. It is the end product of glycolysis.
 - b. It becomes reduced during fermentation.
 - c. It is a precursor of acetyl CoA.
 - d. It is a protein.
 - e. It contains three carbon atoms.
7. The citric acid cycle
 - a. has no connection with the respiratory chain.
 - b. is the same thing as fermentation.
 - c. reduces two NAD⁺ for every glucose processed.
 - d. produces no ATP.
 - e. takes place in the mitochondrion.
8. The respiratory chain
 - a. is located in the mitochondrial matrix.
 - b. includes only peripheral membrane proteins.
 - c. always produces ATP.
 - d. reoxidizes reduced coenzymes.
 - e. operates simultaneously with fermentation.
9. Compared with fermentation, the aerobic pathways of glucose metabolism produce
 - a. more ATP.
 - b. pyruvate.
 - c. fewer protons for pumping in the mitochondria.
 - d. less CO_2 .
 - e. more oxidized coenzymes.
10. Which statement about oxidative phosphorylation is *not* true?
 - a. It forms ATP by the respiratory chain/ATP synthesis.
 - b. It is brought about by chemiosmosis.
 - c. It requires aerobic conditions.
 - d. It takes place in mitochondria.
 - e. Its functions can be served equally well by fermentation.

FOR DISCUSSION

1. Trace the sequence of chemical changes that occurs in mammalian tissue when the oxygen supply is cut off. The first change is that the cytochrome *c* oxidase system becomes totally reduced, because electrons can still flow from cytochrome *c*, but there is no oxygen to accept electrons from cytochrome *c* oxidase. What are the remaining steps?
2. Some cells that use the aerobic pathways of glucose metabolism can also thrive by using fermentation under anaerobic conditions. Given the lower yield of ATP (per molecule

of glucose) in fermentation, how can these cells function so efficiently under anaerobic conditions?

3. The drug antimycin A blocks electron transport in mitochondria. Explain what would happen if the experiment on the left in Figure 9.10 were repeated in the presence of this drug.
4. You eat a burger that contains polysaccharides, proteins, and lipids. Using what you know of the integration of biochemical pathways, explain how the amino acids in the proteins and the glucose in the polysaccharides can end up as fats.

ADDITIONAL INVESTIGATION

A protein in the fat of newborns uncouples the synthesis of ATP from electron transport and instead generates heat. How would

you investigate the hypothesis that this uncoupling protein adds a second proton channel to the mitochondrial membrane?

WORKING WITH DATA (GO TO yourBioPortal.com)

Two Experiments Demonstrate the Chemiosmotic Mechanism

In this real-life exercise, you will examine the background and data from the original research paper by Jagendorf and Uribe in which they showed that an artificially induced H⁺ gradient

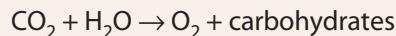
could drive ATP synthesis (Figure 9.10). You will see how they measured ATP by two different methods, and what control experiments they performed to confirm their interpretation.

10

Photosynthesis: Energy from Sunlight

Photosynthesis and global climate change

If all the carbohydrates produced by photosynthesis in a year were in the form of sugar cubes, there would be 300 quadrillion of them. Lined up, these cubes would extend from Earth to Pluto—a lot of photosynthesis! As you may have learned from previous courses, photosynthetic organisms use atmospheric carbon dioxide (CO_2) to produce carbohydrates. The simplified equation says it all:



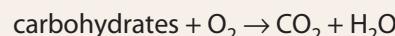
Given the role of CO_2 , how will photosynthesis change with increasing levels of atmospheric CO_2 ? Over the past 200 years, the concentration of atmospheric CO_2 has increased—from 280 parts per million (ppm) in 1800 to 386 ppm in 2008. This increase is correlated with industrialization and the accompanying use of fossil fuels such as coal and oil, which release CO_2 into the atmosphere when they

are burned. The Intergovernmental Panel of Climate Change, sponsored by the United Nations, estimates that atmospheric CO_2 will continue to rise over the next century.

Carbon dioxide is a “greenhouse gas” that traps heat in the atmosphere, and the rising CO_2 level is predicted to result in global climate change. Policy makers concerned about climate change are asking plant biologists to answer two questions about the rise in CO_2 : will it lead to increased photosynthesis, and if so, will it lead to increased plant growth? To answer these questions, scientists initially measured the rate of photosynthesis of plants grown in greenhouses with elevated concentrations of CO_2 . The results were surprising: at first, the rate of photosynthesis went up, but then it returned to near normal as the plants adapted to the higher CO_2 levels.

To determine how plants might respond under more realistic conditions, scientists developed a way to expose plants to high levels of CO_2 in the field. Free-air concentration enrichment (FACE) involves the use of rings of pipes that release CO_2 to the air surrounding plants in fields or forests. Wind speed and direction are monitored by a computer, which constantly controls which pipes release CO_2 . Data from these experiments confirm that photosynthetic rates increase as the concentration of CO_2 rises—although generally the increase is not as high as that seen initially in the greenhouse experiments. Nevertheless, these measurements indicate that as atmospheric CO_2 rises globally, there will be an increase in photosynthesis.

Will this increase in photosynthesis result in an increase in plant growth? Keep in mind that plants, like all organisms, use carbohydrates as an energy source. They perform cellular respiration with the general equation:



Primary Producers Covering less than 2 percent of Earth's surface, rainforests are photosynthetic dynamos. They may act as a “sink” for increasing atmospheric CO_2 .





FACE Free-air carbon dioxide enrichment uses pipes to release CO₂ around plants in the field, to estimate the effects of rising atmospheric CO₂ on photosynthesis and plant growth.

The challenge facing plant biologists is to determine the balance between photosynthesis and respiration and how this affects the rate of plant growth. The FACE experiments indicate that crop yields increase under higher CO₂ concentrations, suggesting that the overall increase in photosynthesis is greater than the increase in respiration. But climate change alters rainfall patterns as well as temperatures. These changes affect where plants grow, and could shift the balance between plant growth and cellular respiration.

As with much in science, the initial questions at first appeared amenable to simple answers. Instead, they led to more questions, and more data are needed. An understanding of the processes of photosynthesis, described in this chapter, provides us with a foundation for asking and answering these urgent questions about climate change and its effects on our world.

IN THIS CHAPTER we begin with a consideration of light energy, and move on to describe how photosynthesis converts light energy into chemical energy, in the form of reduced electron carriers and ATP. Then, we show how these two sources of chemical energy are used to drive the synthesis of carbohydrates from CO₂. Finally, we describe how these processes relate to plant metabolism and growth.

CHAPTER OUTLINE

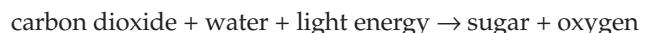
- 10.1** What Is Photosynthesis?
- 10.2** How Does Photosynthesis Convert Light Energy into Chemical Energy?
- 10.3** How Is Chemical Energy Used to Synthesize Carbohydrates?
- 10.4** How Do Plants Adapt to the Inefficiencies of Photosynthesis?
- 10.5** How Does Photosynthesis Interact with Other Pathways?

10.1 What Is Photosynthesis?

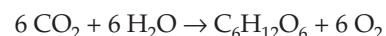
Photosynthesis (literally, “synthesis from light”) is a metabolic process by which the energy of sunlight is captured and used to convert carbon dioxide (CO₂) and water (H₂O) into carbohydrates (which we will represent as a six-carbon sugar, C₆H₁₂O₆) and oxygen gas (O₂). By early in the nineteenth century, scientists had grasped these broad outlines of photosynthesis and had established several facts about the way the process works:

- The water for photosynthesis in land plants comes primarily from the soil, and must travel from the roots to the leaves.
- Plants take in carbon dioxide, producing carbohydrates (sugars) for growth, and plants release O₂ (**Figure 10.1**).
- Light is absolutely necessary for the production of oxygen and sugars.

By 1804, scientists had summarized photosynthesis as follows:



In molecular terms, this equation seems to be the reverse of the overall equation for cellular respiration (see Section 9.1). More precisely, photosynthesis can be written as:

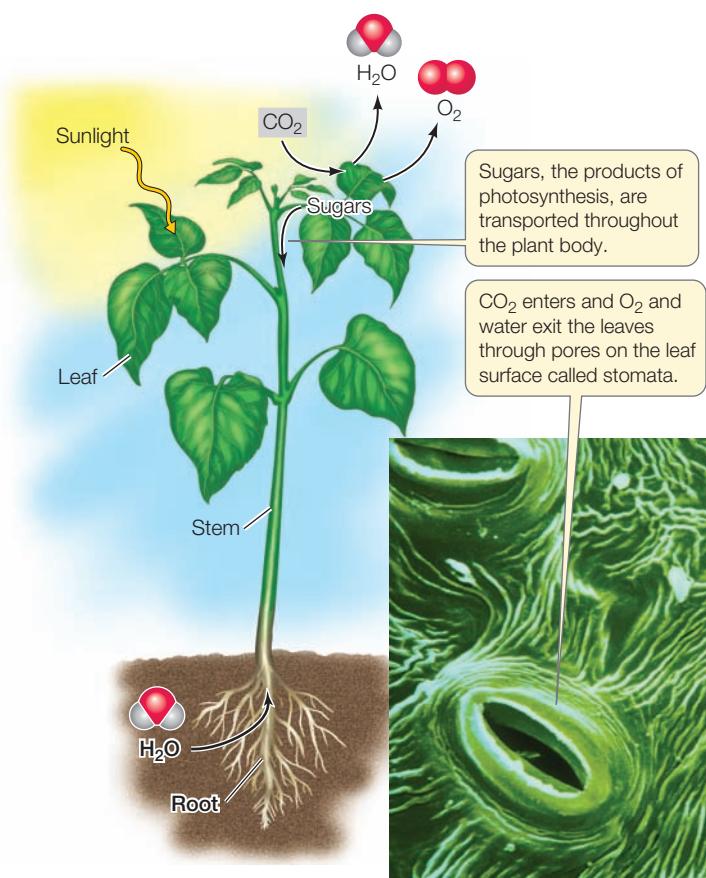


While this equation and the one for cellular respiration (given in the chapter opening story) are essentially correct, they are too general for a real understanding of the processes involved. A number of questions arise: What are the precise chemical reactions of photosynthesis? What role does light play in these reactions? How do carbons become linked to form carbohydrates? What carbohydrates are formed? And where does the oxygen gas come from: CO₂ or H₂O?

Experiments with isotopes show that in photosynthesis O₂ comes from H₂O

In 1941 Samuel Ruben and Martin Kamen, at the University of California, Berkeley, performed experiments using the isotopes ¹⁸O and ¹⁶O to identify the source of the O₂ produced during photosynthesis (**Figure 10.2**). Their results showed that all the oxygen gas produced during photosynthesis comes from water, as is reflected in the revised balanced equation:

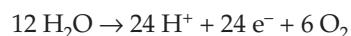




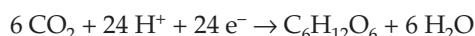
10.1 The Ingredients for Photosynthesis A typical terrestrial plant uses light from the sun, water from the soil, and carbon dioxide from the atmosphere to form organic compounds by photosynthesis.

Water appears on both sides of the equation because it is both used as a reactant (the twelve molecules on the left) and released as a product (the six new ones on the right). This revised equation accounts for all the water molecules needed for all the oxygen gas produced.

The realization that water was the source of photosynthetic O_2 led to an understanding of photosynthesis in terms of *oxidation and reduction*. As we describe in Chapter 9, oxidation-reduction (redox) reactions are coupled: when one molecule becomes oxidized in a reaction, another gets reduced. In this case, oxygen atoms in the reduced state in H_2O get oxidized to O_2 :



while carbon atoms in the oxidized state in CO_2 get reduced to carbohydrate, with the simultaneous production of water:



Adding these two equations (chemistry students will recognize them as *half-cell reactions*) gives the overall equation shown above. As you will see, there is an intermediary carrier of the H^+ and electrons between these two processes—the redox coenzyme, nicotinamide adenine dinucleotide phosphate ($NADP^+$).

Photosynthesis involves two pathways

The equations above summarize the overall process of photosynthesis, but not the stages by which it is completed. Like gly-

colysis and the other metabolic pathways that harvest energy in cells, photosynthesis is a process consisting of many reactions. These reactions are commonly divided into two main pathways:

- The **light reactions** convert light energy into chemical energy in the form of ATP and the reduced electron carrier NADPH. This molecule is similar to NADH (see Section 9.1) but with an additional phosphate group attached to the sugar of its adenosine. In general, NADPH acts as a reducing agent in photosynthesis and other anabolic reactions.
- The **light-independent reactions** (carbon-fixation reactions) do not use light directly, but instead use ATP, NADPH (*made by the light reactions*), and CO_2 to produce carbohydrate.

INVESTIGATING LIFE

10.2 The Source of the Oxygen Produced by Photosynthesis

Although it was clear that O_2 was made during photosynthesis, its molecular source was not known. Two possibilities were the reactants, CO_2 and H_2O . In two separate experiments, Samuel Ruben and Martin Kamen labeled the oxygen in these molecules with the isotope ^{18}O , then tested the O_2 produced by a green plant to find out which molecule contributed the oxygen.

HYPOTHESIS The oxygen released by photosynthesis comes from water rather than CO_2 .

METHOD

Give plants isotope-labeled water and unlabeled CO_2 .

Experiment 1

$H_2^{18}O$, CO_2



Experiment 2

H_2O , $C^{18}O_2$

Give plants isotope-labeled CO_2 and unlabeled water.

RESULTS

The oxygen released is labeled.

$^{18}O_2$

The oxygen released is unlabeled.

CONCLUSION

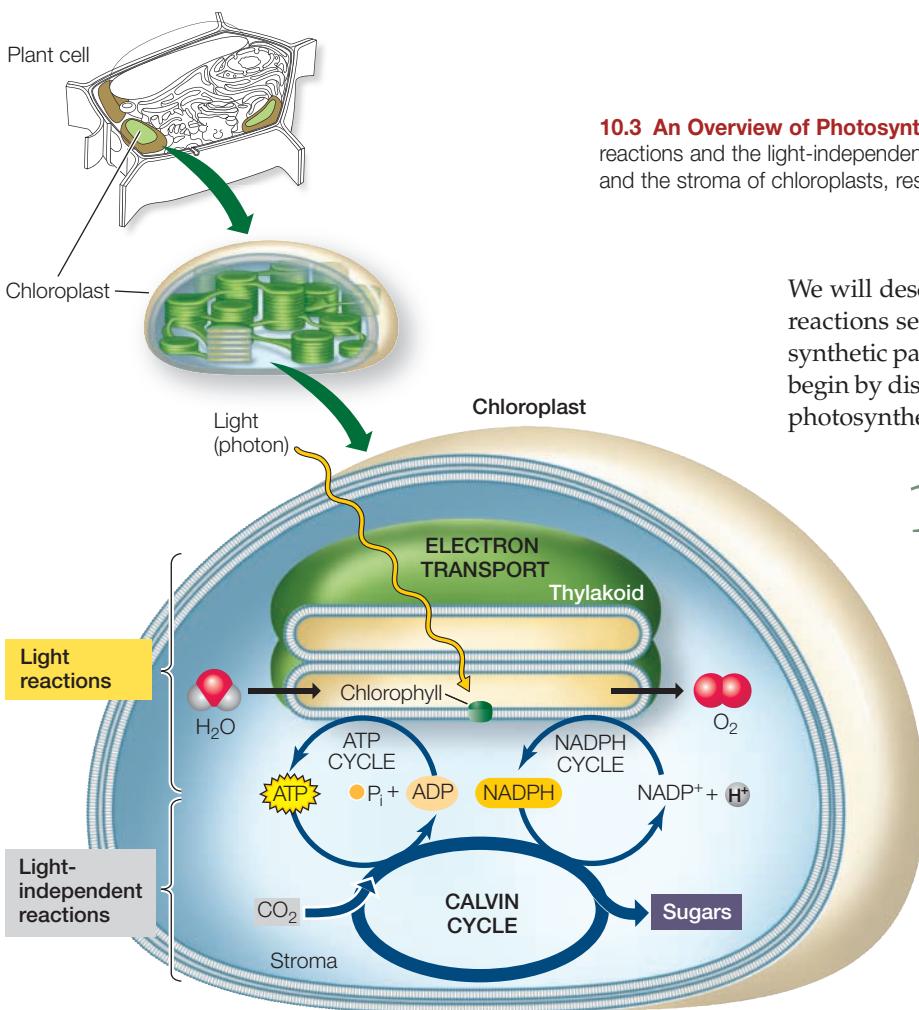
Water is the source of the O_2 produced by photosynthesis.

FURTHER INVESTIGATION: How would you test for the source of oxygen atoms in the carbohydrates made by photosynthesis?

Go to yourBioPortal.com for original citations, discussions, and relevant links for all INVESTIGATING LIFE figures.

yourBioPortal.com

GO TO Animated Tutorial 10.1 • The Source of the Oxygen Produced by Photosynthesis



10.3 An Overview of Photosynthesis Photosynthesis consists of two pathways: the light reactions and the light-independent reactions. These reactions take place in the thylakoids and the stroma of chloroplasts, respectively.

We will describe the light reactions and the light-independent reactions separately and in detail. But since these two photosynthetic pathways are powered by the energy of sunlight, let's begin by discussing the physical nature of light and the specific photosynthetic molecules that capture its energy.

10.2 How Does Photosynthesis Convert Light Energy into Chemical Energy?

Light is a form of energy, and it can be converted to other forms of energy such as heat or chemical energy. Our focus here will be on light as the source of energy to drive the formation of ATP (from ADP and P_i) and NADPH (from NADP⁺ and H⁺).

Light is a form of energy with dual properties

Light is a form of **electromagnetic radiation**. It is propagated in waves, and the amount of energy in light is inversely proportional to its **wavelength**—the shorter the wavelength, the greater the energy. The visible portion of the electromagnetic spectrum (**Figure 10.4**) encompasses a wide range of wavelengths and energy levels. In addition to traveling in waves, light also behaves as particles, called **photons**, which have no mass. In plants and other photosynthetic organisms, receptive molecules absorb photons in order to harvest their energy for biological processes. Because these receptive molecules absorb only specific wavelengths of light, the photons must have the correct amount of energy—they must be of the appropriate wavelength.

Molecules become excited when they absorb photons

When a photon meets a molecule, one of three things can happen:

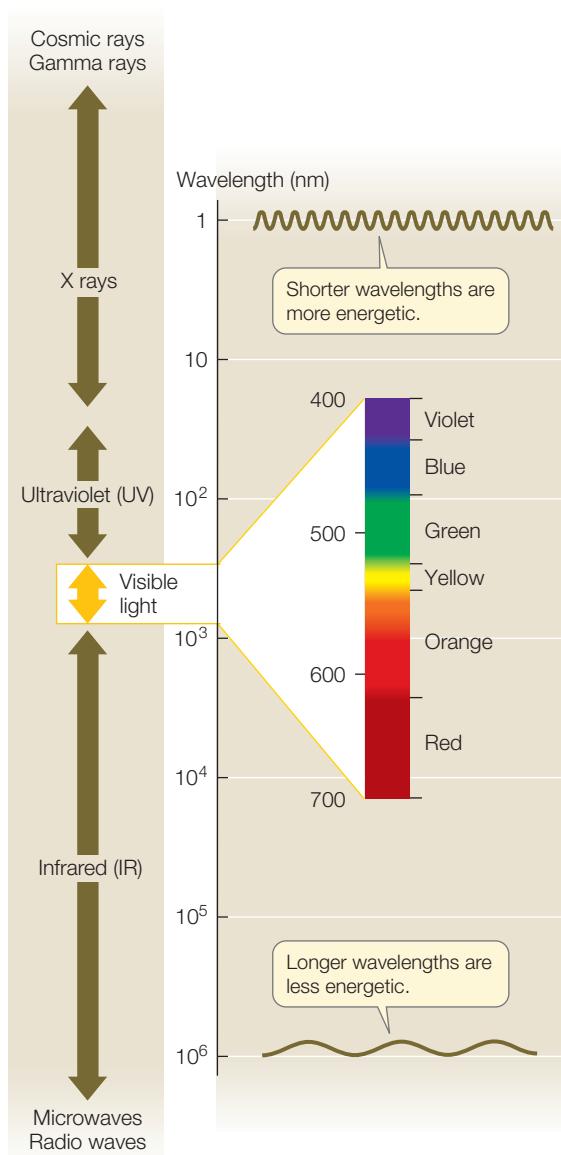
- The photon may bounce off the molecule—it may be scattered or reflected.
- The photon may pass through the molecule—it may be transmitted.
- The photon may be absorbed by the molecule, adding energy to the molecule.

Neither of the first two outcomes causes any change in the molecule. However, in the case of absorption, the photon disappears and its energy is absorbed by the molecule. The photon's energy cannot disappear, because according to the first law of thermodynamics, energy is neither created nor destroyed. When the molecule acquires the energy of the photon it is raised from a ground state (with lower energy) to an excited state (with higher energy) (**Figure 10.5A**).

10.1 RECAP

The light reactions of photosynthesis convert light energy into chemical energy. The light-independent reactions use that chemical energy to reduce CO₂ to carbohydrates.

- What is the experimental evidence that water is the source of the O₂ produced during photosynthesis? See pp. 190–191 and **Figure 10.2**
- What is the relationship between the light reactions and the light-independent reactions of photosynthesis? See pp. 191–192 and **Figure 10.3**



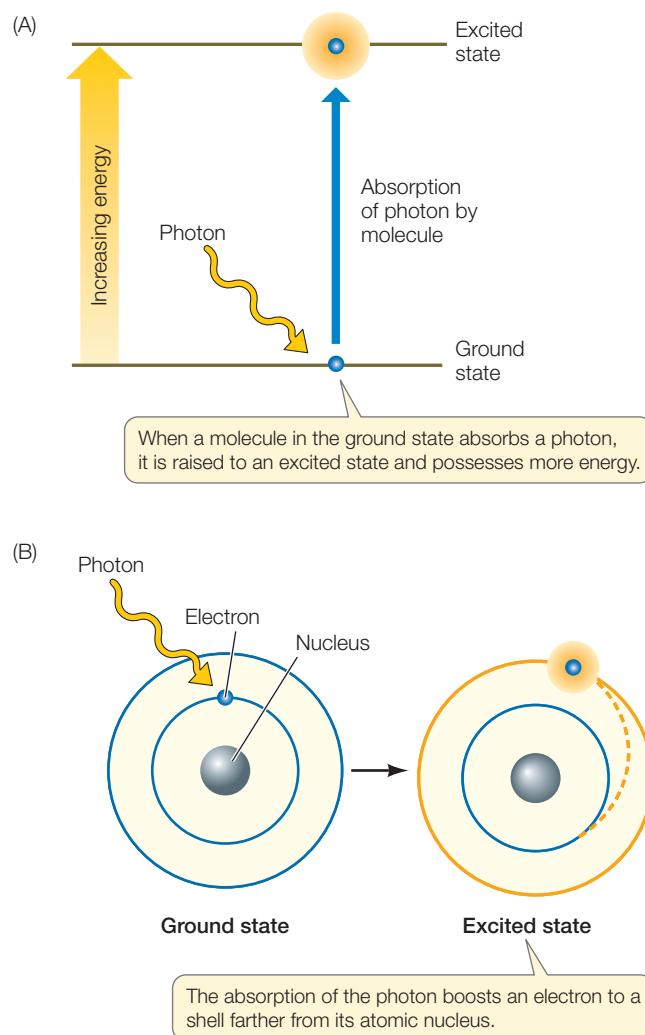
10.4 The Electromagnetic Spectrum The portion of the electromagnetic spectrum that is visible to humans as light is shown in detail at the right.

The difference in free energy between the molecule's excited state and its ground state is approximately equal to the free energy of the absorbed photon (a small amount of energy is lost to entropy). The increase in energy boosts one of the electrons within the molecule into a shell farther from its nucleus; this electron is now held less firmly (Figure 10.5B), making the molecule unstable and more chemically reactive.

Absorbed wavelengths correlate with biological activity

The specific wavelengths absorbed by a particular molecule are characteristic of that type of molecule. Molecules that absorb wavelengths in the visible spectrum are called **pigments**.

When a beam of white light (containing all the wavelengths of visible light) falls on a pigment, certain wavelengths are absorbed. The remaining wavelengths, which are scattered or trans-

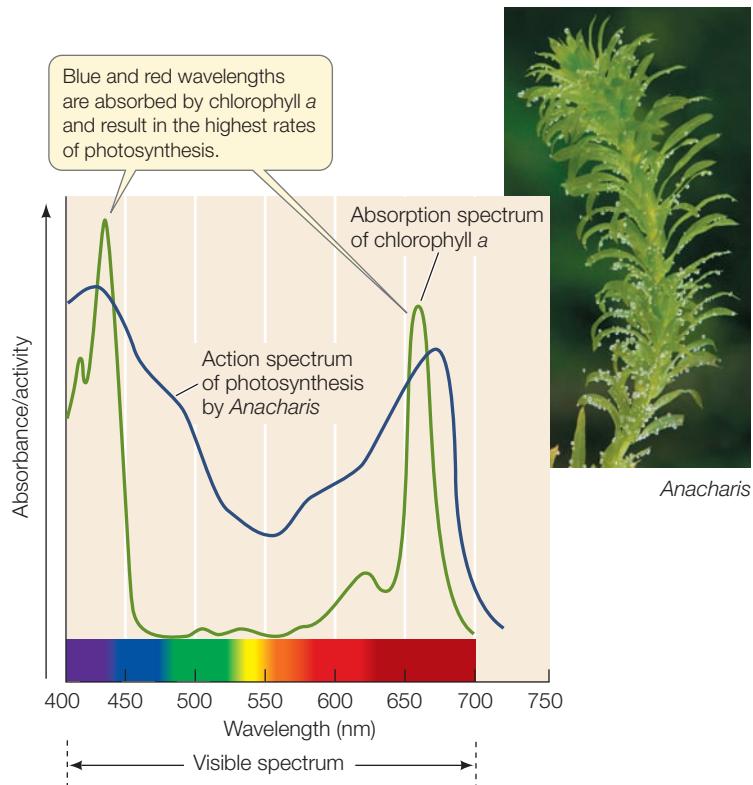


10.5 Exciting a Molecule (A) When a molecule absorbs the energy of a photon, it is raised from a ground state to an excited state. (B) In the excited state, an electron is boosted to a shell more distant from the atomic nucleus, where it is held less firmly.

mitted, make the pigment appear to us as colored. For example, if a pigment absorbs both blue and red light (as does chlorophyll) what we see is the remaining light, which is primarily green. If we plot light absorbed by a purified pigment against wavelength, the result is an **absorption spectrum** for that pigment.

In contrast to the absorption spectrum, an **action spectrum** is a plot of the *biological activity* of an organism as a function of the wavelengths of light to which it is exposed. The experimental determination of an action spectrum might be performed as follows:

1. Place a plant (a water plant with thin leaves is convenient) in a closed container.
2. Expose the plant to light of a certain wavelength for a period of time.
3. Measure photosynthesis by the amount of O_2 released.
4. Repeat with light of other wavelengths.



10.6 Absorption and Action Spectra The absorption spectrum of the purified pigment chlorophyll *a* from the aquatic plant *Anacharis* is similar to the action spectrum obtained when different wavelengths of light are shone on the intact plant and the rate of photosynthesis is measured. In the thicker leaves of land plants, the action spectra show less of a dip in the green region (500–650 nm).

Figure 10.6 shows the absorption spectrum of the pigment chlorophyll *a*, which was isolated from the leaves of *Anacharis*, a common aquarium plant. Also shown is the action spectrum for photosynthetic activity by the same plant. A comparison of the two spectra shows that the wavelengths at which photosynthesis is highest are the same wavelengths at which chlorophyll *a* absorbs light.

Several pigments absorb energy for photosynthesis

The light energy used for photosynthesis is not absorbed by just one type of pigment. Instead, several different pigments with different absorption spectra collect the energy that is eventually used for photosynthesis. In photosynthetic organisms as diverse as green algae, protists, and bacteria, these pigments include chlorophylls, carotenoids, and phycobilins.

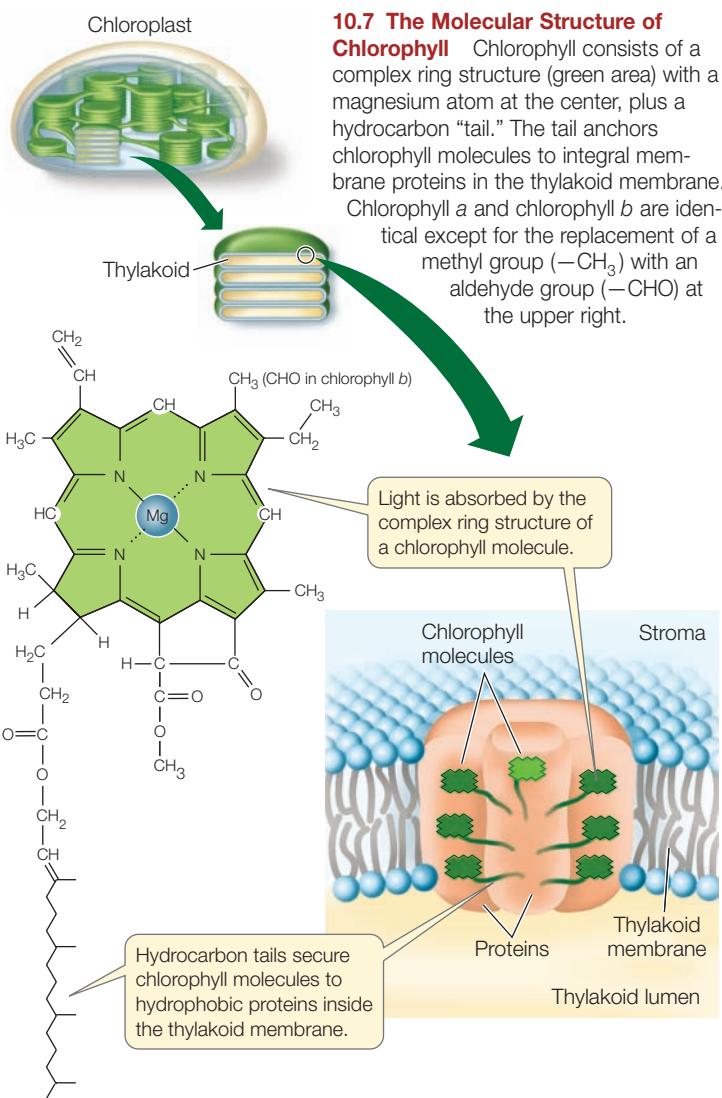
CHLOROPHYLLS In plants, two chlorophylls are responsible for absorbing the light energy that is used to drive the light reactions: **chlorophyll *a*** and **chlorophyll *b***. These two molecules differ only slightly in their molecular structures. Both have a complex ring structure similar to that of the heme group of hemoglobin (**Figure 10.7**). In the center of the chlorophyll ring is a magnesium atom. Attached at a peripheral location on the ring is a long hydrocarbon “tail,” which anchors the chlorophyll molecule to in-

tegral proteins in the thylakoid membrane of the chloroplast. (See Figure 5.13 to review the anatomy of a chloroplast.)

ACCESSORY PIGMENTS We saw in Figure 10.6 that chlorophyll absorbs blue and red light, which are near the two ends of the visible spectrum. Thus, if only chlorophyll were active in photosynthesis, much of the visible spectrum would go unused. This appears to be the case in higher plants. But lower plants (such as algae) and cyanobacteria possess accessory pigments, which absorb photons intermediate in energy between the red and the blue wavelengths and then transfer a portion of that energy to the chlorophylls. Among these accessory pigments are **carotenoids** such as β -carotene (see Figure 3.21), which absorb photons in the blue and blue-green wavelengths and appear deep yellow. The **phycobilins**, which are found in red algae and in cyanobacteria, absorb various yellow-green, yellow, and orange wavelengths.

Light absorption results in photochemical change

Any pigment molecule can become excited when its absorption spectrum matches the energies of incoming photons. After a



pigment molecule absorbs a photon and enters an excited state (see Figure 10.5), there are several alternative fates for the absorbed energy:

- It can be released as heat and/or light.
- It may be rapidly transferred to a neighboring pigment molecule.
- It can be used as free energy to drive a chemical reaction.

When the excited molecule gives up the absorbed energy it returns to the ground state.

Sometimes the absorbed energy is given off as heat and light, in a process called *fluorescence*. Because some of the energy of the original absorbed photon is lost as heat, the photon that is released as fluorescence has less energy and a longer wavelength than the absorbed light. When there is fluorescence, there are no permanent chemical changes made or biological functions performed—no chemical work is done.

On the other hand, the excited pigment molecule may pass the absorbed energy along to another molecule—provided that the target molecule is very near, has the right orientation, and has the appropriate structure to receive the energy. This is what happens in photosynthesis.

The pigments in photosynthetic organisms are arranged into energy-absorbing **antenna systems**, also called *light-harvesting complexes*. These form part of a large multi-protein complex called a **photosystem**. The photosystem spans the thylakoid membrane, and consists of multiple antenna systems, with their associated pigment molecules, all surrounding a **reaction center**. The pigment molecules in the antenna systems are packed together in such a way that the excitation energy from an absorbed photon can be passed along from one pigment molecule to another (Figure 10.8). Excitation energy moves from pigments that absorb shorter wavelengths (higher energy) to pigments that absorb

longer wavelengths (lower energy). Thus the excitation ends up in the pigment molecules that absorb the longest wavelengths. These pigment molecules are in the reaction center of the photosystem, and form special associations with the photosystem proteins (see Figure 10.8). The ratio of antenna pigments to reaction center pigments can be quite high (over 300:1).

The reaction center converts the absorbed light energy into chemical energy. A pigment molecule in the reaction center absorbs sufficient energy that it actually gives up its excited electron (is chemically oxidized) and becomes positively charged. In plants, the reaction center contains a pair of chlorophyll *a* molecules. There are many other chlorophyll *a* molecules in the antenna systems, but because of their interactions with antenna proteins, all of them absorb light at shorter wavelengths than the pair in the reaction center.

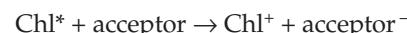
Excited chlorophylls in the reaction center act as electron donors

Chlorophyll has two vital roles in photosynthesis:

- It absorbs light energy and transforms it into excited electrons.
- It transfers those electrons to other molecules, initiating chemical changes.

We have dealt with the first role; now we turn to the second.

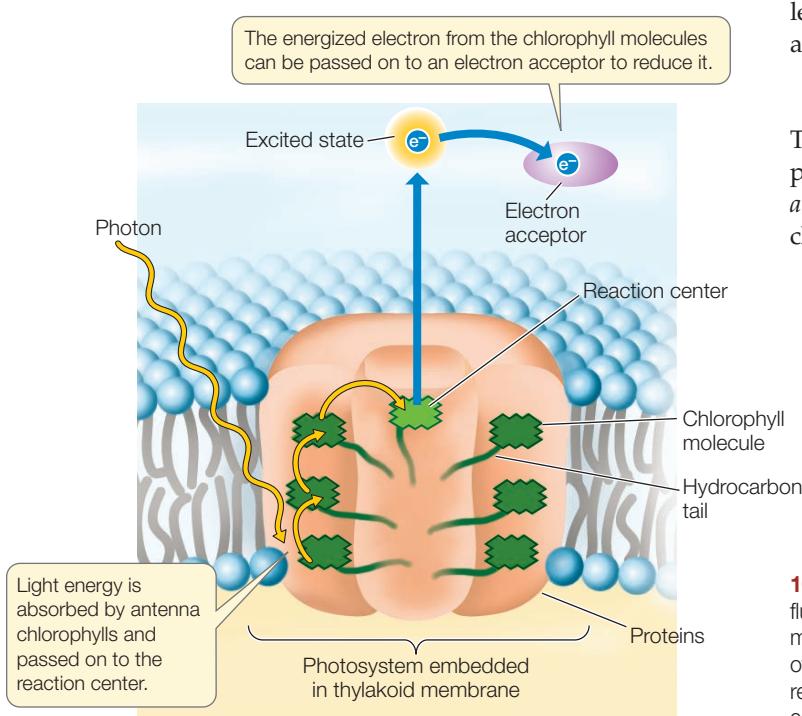
Photosynthesis harvests chemical energy by using the excited chlorophyll molecules in the reaction center as electron donors (reducing agents) to reduce a stable electron acceptor (see Figure 10.8). Ground-state chlorophyll (symbolized by Chl) is not much of a reducing agent, but excited chlorophyll (Chl*) is a good one. This is because in the excited molecule, one of the electrons has moved to a shell that is farther away from the nucleus than the shell it normally occupies. This electron is held less tightly than in the normal state, and it can be transferred in a redox reaction to an electron acceptor (an oxidizing agent):



This, then, is the first consequence of light absorption by chlorophyll: a reaction center chlorophyll (Chl*) loses its excited electron in a redox reaction and becomes Chl⁺ (because it gives up a negative charge—it gets oxidized).

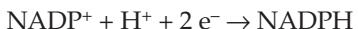
Reduction leads to electron transport

The electron acceptor that is reduced by Chl* is the first in a chain of electron carriers in the thylakoid membrane that participate in a process termed *electron transport*. This energetically “downhill” series of reductions and oxidations is similar to what occurs in the respiratory chain of



10.8 Energy Transfer and Electron Transport Rather than being lost as fluorescence, energy from a photon may be transferred from one pigment molecule to another. In a photosystem, energy is transferred through a series of molecules to one or more pigment molecules in the reaction center. If a reaction center molecule becomes sufficiently excited, it will give up its excited electron to an electron acceptor.

mitochondria (see Section 9.3). The final electron acceptor is NADP⁺ (nicotinamide adenine dinucleotide phosphate), which gets reduced:



The energy-rich NADPH is a stable, reduced coenzyme.

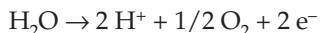
There are two different systems of electron transport in photosynthesis:

- **Noncyclic electron transport** produces NADPH and ATP. Essentially, the excited electron is “lost” from chlorophyll and the transport process ends up with a reduced coenzyme.
- **Cyclic electron transport** produces only ATP. Essentially, the transport process ends up with the excited electron returning to chlorophyll, after giving up energy to make ATP.

We'll consider these two systems before describing the production of ATP from ADP and P_i.

Noncyclic electron transport produces ATP and NADPH

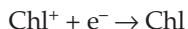
In noncyclic electron transport, light energy is used to oxidize water, forming O₂, H⁺, and electrons. In quantitative terms this would be



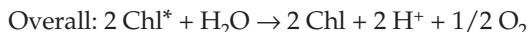
We saw above that a key reaction in photosynthesis occurs when chlorophyll that is excited by absorbing light (Chl*) gives up its excited electron, becoming oxidized:



Because it lacks an electron, Chl⁺ is very unstable; it has a very strong tendency to “grab” an electron from another molecule to replenish the one it lost:



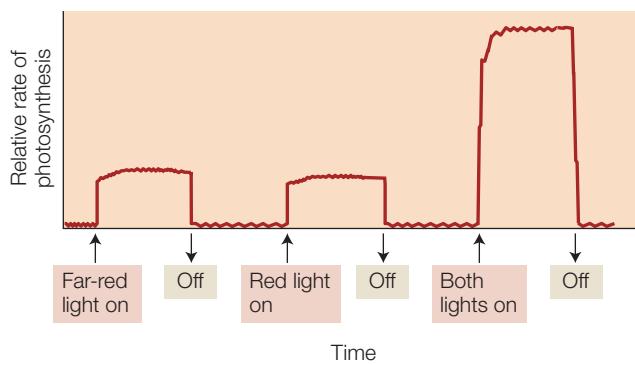
So in chemical terms, Chl⁺ is a strong oxidizing agent. The replenishing electrons come from water, splitting the H–O–H bonds.



Notice that this is a more precise description of what Ruben and Kamen had found, namely that the source of O₂ in photosynthesis is H₂O (see Figure 10.2).

The electrons are passed from chlorophyll to NADP⁺ through a chain of electron carriers in the thylakoid membrane. These redox reactions are exergonic, and some of the released free energy is ultimately used to form ATP by *chemiosmosis* (see p. 180).

TWO PHOTOSYSTEMS ARE REQUIRED Noncyclic electron transport requires the participation of two different photosystems in the thylakoid membrane. What is the evidence of the existence of these two photosystems? In 1957, Robert Emerson at the University of Illinois shone light of various wavelengths onto cells of *Chlorella*, a freshwater protist. Both red light (wavelength 680 nm) and far-red light (700 nm) resulted in modest rates of photosynthesis, as measured by O₂ production. But when the two lights



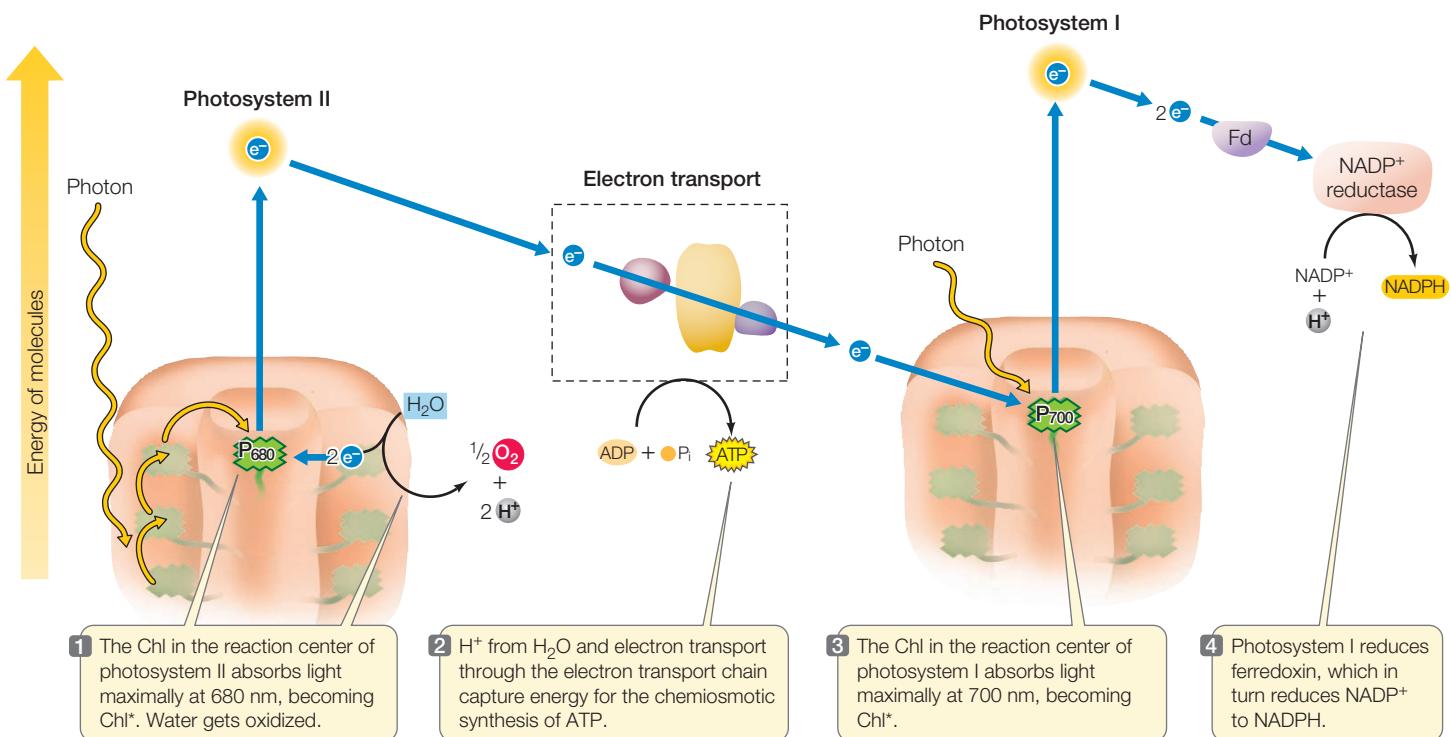
10.9 Two Photosystems The absorption and action spectra for chlorophyll and photosynthesis indicated that the rate of photosynthesis would increase in red light. Robert Emerson shone red (660 nm) and far-red (>700 nm) light both separately and together on algal cells to look for cooperative effects.

were combined, the rate of photosynthesis was much greater than the rates under either red light or far red light. In fact it was greater than the two rates added together. This phenomenon was termed photo enhancement (Figure 10.9). A few years later, photo enhancement was explained by the existence of *not one but two reaction centers*, which act together to enhance photosynthesis.

- **Photosystem I** uses light energy to pass an excited electron to NADP⁺, reducing it to NADPH.
- **Photosystem II** uses light energy to oxidize water molecules, producing electrons, protons (H⁺), and O₂.

The reaction center for photosystem I contains a pair of chlorophyll *a* molecules called P₇₀₀ because it can best absorb light with a wavelength of 700 nm. Similarly, the pair of chlorophyll *a* molecules in the photosystem II reaction center is called P₆₈₀ because it absorbs light maximally at 680 nm. Thus photosystem II requires photons that are somewhat more energetic (i.e., have shorter wavelengths) than those required by photosystem I. To keep noncyclic electron transport going, both photosystems must be constantly absorbing light, thereby boosting electrons to higher shells from which they may be captured by specific electron acceptors. A model for the way photosystems I and II interact and complement each other is called the “Z scheme,” because when the path of the electrons is placed along an axis of rising energy level, it resembles a sideways letter Z (Figure 10.10).

ELECTRON TRANSPORT: THE Z SCHEME In the Z scheme model, which describes the reactions of noncyclic electron transport from water to NADP⁺, photosystem II comes before photosystem I. When photosystem II absorbs photons, electrons pass from P₆₈₀ to the primary electron acceptor and P₆₈₀* is oxidized to P₆₈₀⁺. Then an electron from the oxidation of water is passed to P₆₈₀⁺, reducing it to P₆₈₀ once again, so that it can receive more energy from neighboring chlorophyll molecules in the antenna systems. The electrons from photosystem II pass through a series of transfer reactions, one of which is directly responsible for the physical movement of protons from the stroma (the matrix outside the thylakoids) across the thylakoid membrane and into the lumen (see Figure 10.12). In addition to these protons, the protons derived from the splitting of water are deposited into the thylakoid lumen. Furthermore, protons in the stroma are



consumed during the reduction of NADP⁺, and together these reactions create a proton gradient across the thylakoid membrane, which provides the energy for ATP synthesis.

In photosystem I, the P₇₀₀ molecules in the reaction center become excited to P₇₀₀^{*}, leading to the reduction of an electron carrier called ferredoxin (Fd) and the production of P₇₀₀⁺. P₇₀₀⁺ returns to the reduced state by accepting electrons passed through the electron transport system from photosystem II. Having identified the role of the electrons produced by photosystem II, we can now ask, "What is the role of the electrons transferred to Fd from photosystem I?" These electrons are used in the last step of noncyclic electron transport, in which two electrons and a proton are used to reduce a molecule of NADP⁺ to NADPH.

In summary:

- Noncyclic electron transport extracts electrons from water and passes them ultimately to NADPH, utilizing energy absorbed by photosystems I and II, and resulting in ATP synthesis.
- Noncyclic electron transport yields NADPH, ATP, and O₂.

Cyclic electron transport produces ATP but no NADPH

Noncyclic electron transport results in the production of ATP and NADPH. However, as we will see, the light-independent reactions of photosynthesis require more ATP than NADPH + H⁺. If only the noncyclic pathway is operating, there is the possibility that there will not be enough ATP formed. **Cyclic electron transport** makes up for the imbalance. This pathway, which produces only ATP, is called *cyclic* because an electron passed from an excited chlorophyll molecule at the outset cycles back to the same chlorophyll molecule at the end of the chain of reactions (Figure 10.11).

10.10 Noncyclic Electron Transport Uses Two Photosystems

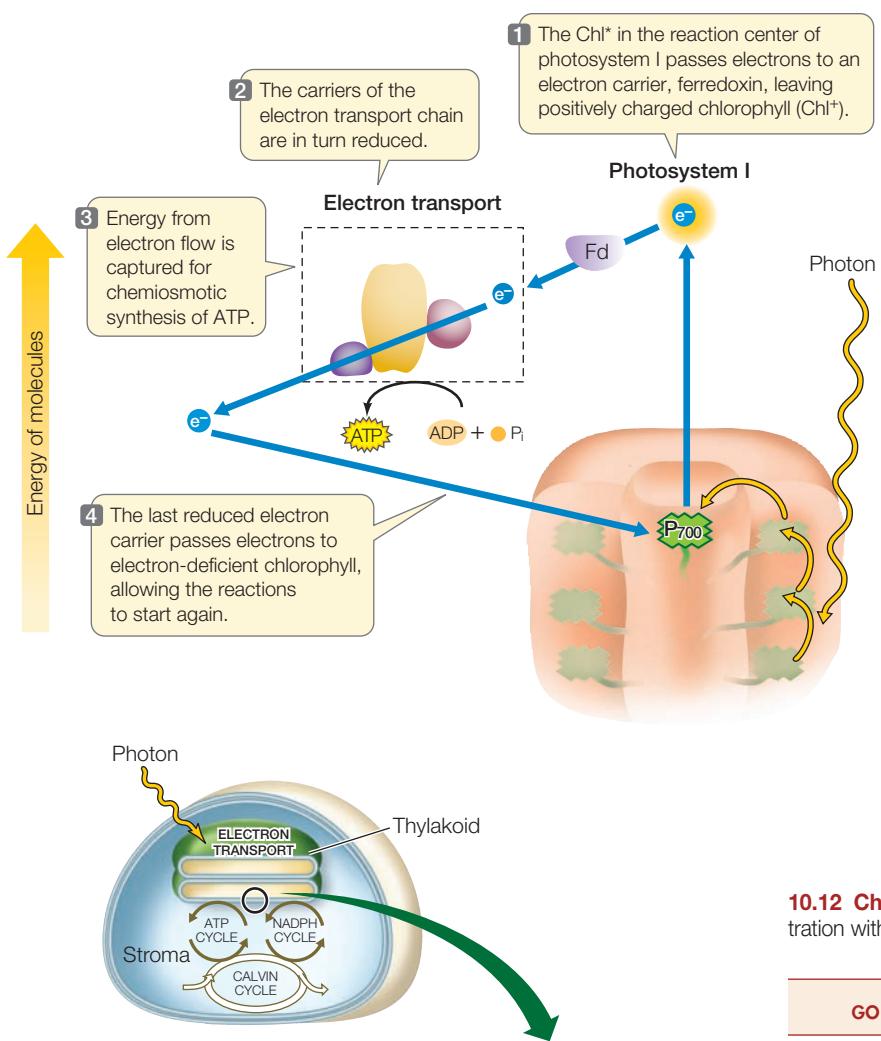
Absorption of light energy by chlorophyll molecules in the reaction centers of photosystems I and II allows them to pass electrons into a series of redox reactions. The term "Z scheme" describes the path (blue arrows) of electrons as they travel through the two photosystems. On this scheme the vertical positions represent the energy levels of the molecules in the electron transport system.

Cyclic electron transport begins and ends in photosystem I. A P₇₀₀ chlorophyll molecule in the reaction center absorbs a photon and enters the excited state, P₇₀₀^{*}. The excited electron is passed from P₇₀₀^{*} to a primary acceptor, and then to oxidized ferredoxin (Fd_{ox}), producing reduced ferredoxin (Fd_{red}). Fd_{red} passes its added electron to a different oxidizing agent, plastocyanin (PQ, a small organic molecule), resulting in the transfer of two H⁺ from the stroma to the thylakoid lumen. The electron passes from reduced PQ through the electron transport system until it completes its cycle by returning to P₇₀₀⁺, restoring it to its uncharged form, P₇₀₀. This electron transport is carried out by plastocyanin (PC) and cytochromes that are similar to those of the mitochondrial respiratory chain.

By the time the electron from P₇₀₀^{*} travels through the electron transport system and comes back to reduce P₇₀₀⁺, all the energy from the original photon has been released. The released energy is stored in the form of a proton gradient that can be used to produce ATP.

Chemiosmosis is the source of the ATP produced in photophosphorylation

In Chapter 9 we describe the chemiosmotic mechanism for ATP formation in the mitochondrion. A similar mechanism, called **photophosphorylation**, operates in the chloroplast, where electron transport is coupled to the transport of protons (H⁺) across



10.11 Cyclic Electron Transport Traps Light Energy as ATP

Cyclic electron transport produces ATP, but no NADPH.

the thylakoid membrane, resulting in a proton gradient across the membrane (Figure 10.12).

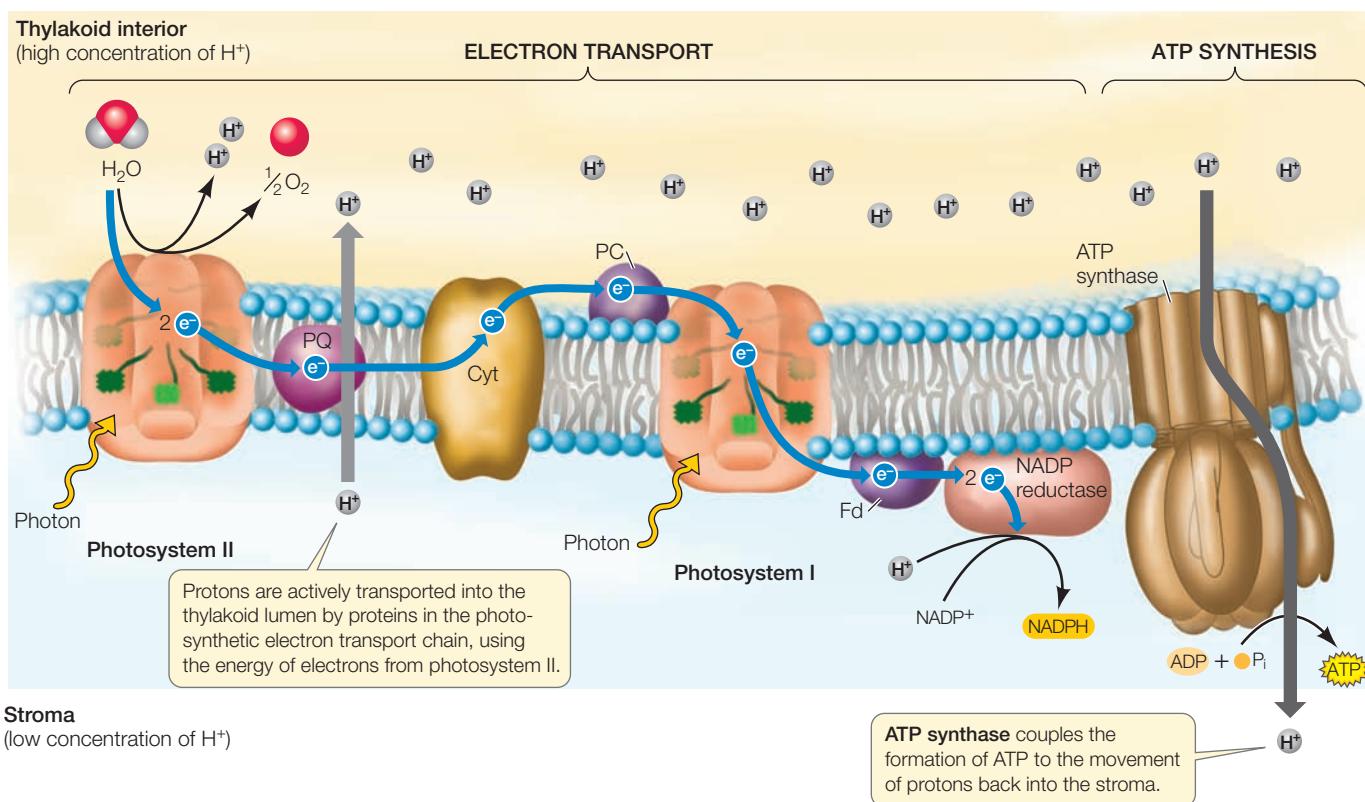
The electron carriers in the thylakoid membrane are oriented so that protons are actively pumped from the stroma into the lumen of the thylakoid. Thus the lumen becomes acidic with respect to the stroma, resulting in an electrochemical gradient across the thylakoid membrane, whose bilayer is not permeable to H^+ . Water oxidation and $NADP^+$ reduction also contribute to this gradient, which drives the movement of protons back out of the thylakoid lumen through specific protein channels in the thylakoid membrane. These channels are enzymes—ATP synthases—that couple the movement of protons to the formation of ATP, as they do in mitochondria (see Figure 9.9). Indeed, chloroplast ATP synthase is about 60 percent identical to human mitochondrial ATP synthase—a remarkable similarity, given that plants and animals had their most recent

10.12 Chloroplasts Form ATP Chemiosmotically

Compare this illustration with Figure 9.9, where a similar process is depicted in mitochondria.

yourBioPortal.com

GO TO Animated Tutorial 10.2 • Photophosphorylation



INVESTIGATING LIFE

common ancestor more than a billion years ago. This is testimony to the evolutionary unity of life.

The mechanisms of the two enzymes are similar, but their orientations differ. In chloroplasts, protons flow through the ATP synthase out of the thylakoid lumen into the stroma (where the ATP is synthesized) but in mitochondria they flow out of the cytosol into the mitochondrial matrix.

10.2 RECAP

Conversion of light energy into chemical energy occurs when pigments absorb photons. Light energy is used to drive a series of protein-associated redox reactions in the thylakoid membranes of the chloroplast.

- How does chlorophyll absorb and transfer light energy? See pp. 194–195 and Figure 10.8
- How are electrons produced in photosystem II and how do they flow to photosystem I? See pp. 196–197 and Figure 10.10
- How does cyclic electron transport in photosystem I result in the production of ATP? See p. 197 and Figure 10.11

We have seen how light energy drives the synthesis of ATP and NADPH in the stroma of chloroplasts. We now turn to the light-independent reactions of photosynthesis, which use energy-rich ATP and NADPH to reduce CO_2 and form carbohydrates.

10.3 How Is Chemical Energy Used to Synthesize Carbohydrates?

Most of the enzymes that catalyze the reactions of CO_2 fixation are dissolved in the stroma of the chloroplast, where those reactions take place. These enzymes use the energy in ATP and NADPH to reduce CO_2 to carbohydrates. Therefore, with some exceptions, CO_2 fixation occurs only in the light, when ATP and NADPH are being generated.

Radioisotope labeling experiments revealed the steps of the Calvin cycle

To identify the reactions by which the carbon from CO_2 ends up in carbohydrates, scientists found a way to label CO_2 so that they could isolate and identify the compounds formed from it during photosynthesis. In the 1950s, Melvin Calvin, Andrew Benson, and their colleagues used radioactively labeled CO_2 in which some of the carbon atoms were the radioisotope ^{14}C rather than the normal ^{12}C . Although ^{14}C emits radiation, its chemical behavior is virtually identical to that of nonradioactive ^{12}C .

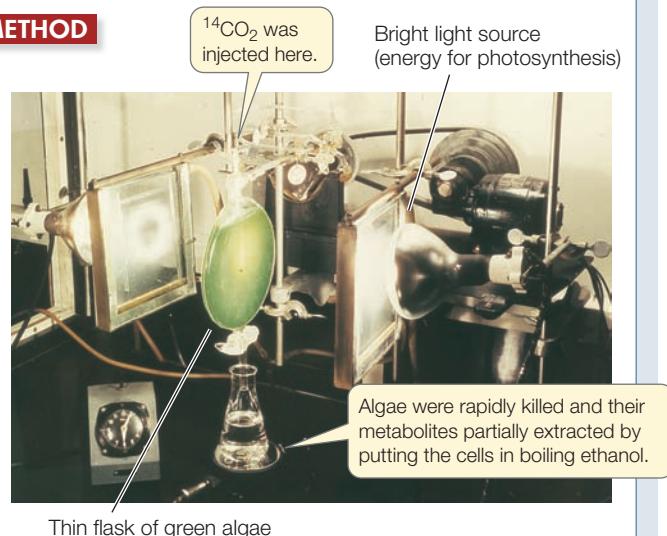
Calvin and his colleagues exposed cultures of the unicellular green alga *Chlorella* to $^{14}\text{CO}_2$ for various lengths of time. Then they rapidly killed the cells and extracted the organic compounds. They separated the different compounds from one another by paper chromatography and exposed the paper to X-ray film (Figure 10.13). When the film was developed, dark

10.13 Tracing the Pathway of CO_2

How is CO_2 incorporated into carbohydrate during photosynthesis? What is the first stable covalent linkage that forms with the carbon of CO_2 ? Short exposures to $^{14}\text{CO}_2$ were used to identify the first compound formed from CO_2 .

HYPOTHESIS The first product of CO_2 fixation is a 3-carbon molecule.

METHOD

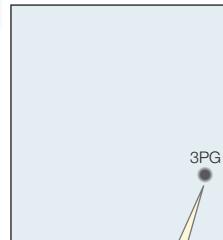


Thin flask of green algae

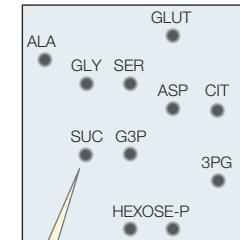
The algal extract was spotted here and run in two directions to separate compounds from one another.

After separation of the compounds, the chromatogram was overlaid with X-ray film.

RESULTS



A chromatogram made after 3 seconds of exposure to $^{14}\text{CO}_2$ shows ^{14}C only in 3PG (3-phosphoglycerate).



A chromatogram made after 30 seconds of exposure to $^{14}\text{CO}_2$ shows ^{14}C in many molecules.

CONCLUSION

The initial product of CO_2 fixation is 3PG. Later, the carbon from CO_2 ends up in many molecules.

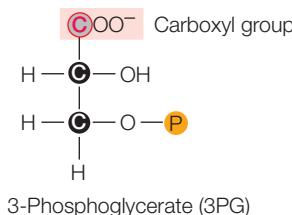
Go to [yourBioPortal.com](#) for original citations, discussions, and relevant links for all INVESTIGATING LIFE figures.

yourBioPortal.com

GO TO Animated Tutorial 10.3 • Tracing the Pathway of CO_2

spots indicated the locations of compounds containing ^{14}C in the paper.

To discover the first compound in the pathway of CO_2 fixation, Calvin and his team exposed the algae to $^{14}\text{CO}_2$ for shorter and shorter periods of time. The 3-second exposure revealed that only one compound was labeled—a 3-carbon sugar phosphate called 3-phosphoglycerate (3PG) (the ^{14}C is shown in red):



With successive exposures longer than 3 seconds, Calvin and his colleagues were able to trace the route of ^{14}C as it moved through a series of compounds, including monosaccharides and amino acids. It turned out that the pathway the ^{14}C moved through was a cycle. In this cycle, the CO_2 initially bonds covalently to a larger five-carbon acceptor molecule, which then breaks into two three-carbon molecules. As the cycle repeats a carbohydrate is produced and the initial CO_2 acceptor is regenerated. This was appropriately named the **Calvin cycle**.

The initial reaction in the Calvin cycle adds the 1-carbon CO_2 to an acceptor molecule, the 5-carbon compound ribulose 1,5-bisphosphate (RuBP). The product is an intermediate 6-carbon compound, which quickly breaks down and forms two molecules of 3PG (Figure 10.14). The intermediate compound is broken down so rapidly that Calvin did not observe radioactive label appearing in it first. But the enzyme that catalyzes its formation, **ribulose bisphosphate carboxylase/oxygenase (rubisco)**, is the most abundant protein in the world! It constitutes up to 50 percent of all the protein in every plant leaf.

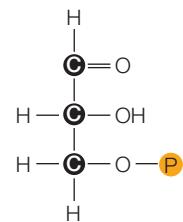
The Calvin cycle is made up of three processes

The Calvin cycle uses the ATP and NADPH made in the light to reduce CO_2 in the stroma to a carbohydrate. Like all biochem-

ical pathways, each reaction is catalyzed by a specific enzyme. The cycle is composed of three distinct processes (Figure 10.15):

- *Fixation* of CO_2 . As we have seen, this reaction is catalyzed by rubisco, and its stable product is 3PG.
- *Reduction* of 3PG to form glyceraldehyde 3-phosphate (G3P). This series of reactions involves a phosphorylation (using the ATP made in the light reactions) and a reduction (using the NADPH made in the light reactions).
- *Regeneration* of the CO_2 acceptor, RuBP. Most of the G3P ends up as ribulose monophosphate (RuMP), and ATP is used to convert this compound into RuBP. So for every “turn” of the cycle, with one CO_2 fixed, the CO_2 acceptor is regenerated.

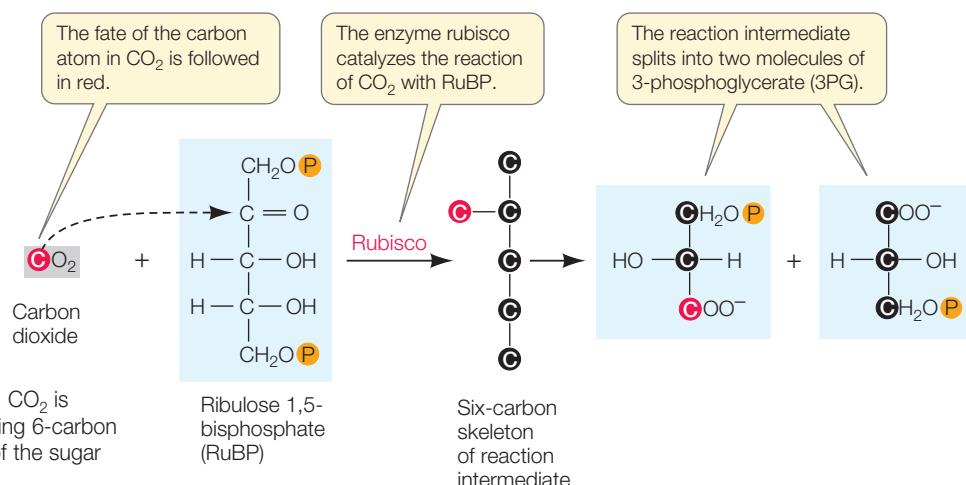
The product of this cycle is **glyceraldehyde 3-phosphate (G3P)**, which is a 3-carbon sugar phosphate, also called triose phosphate:



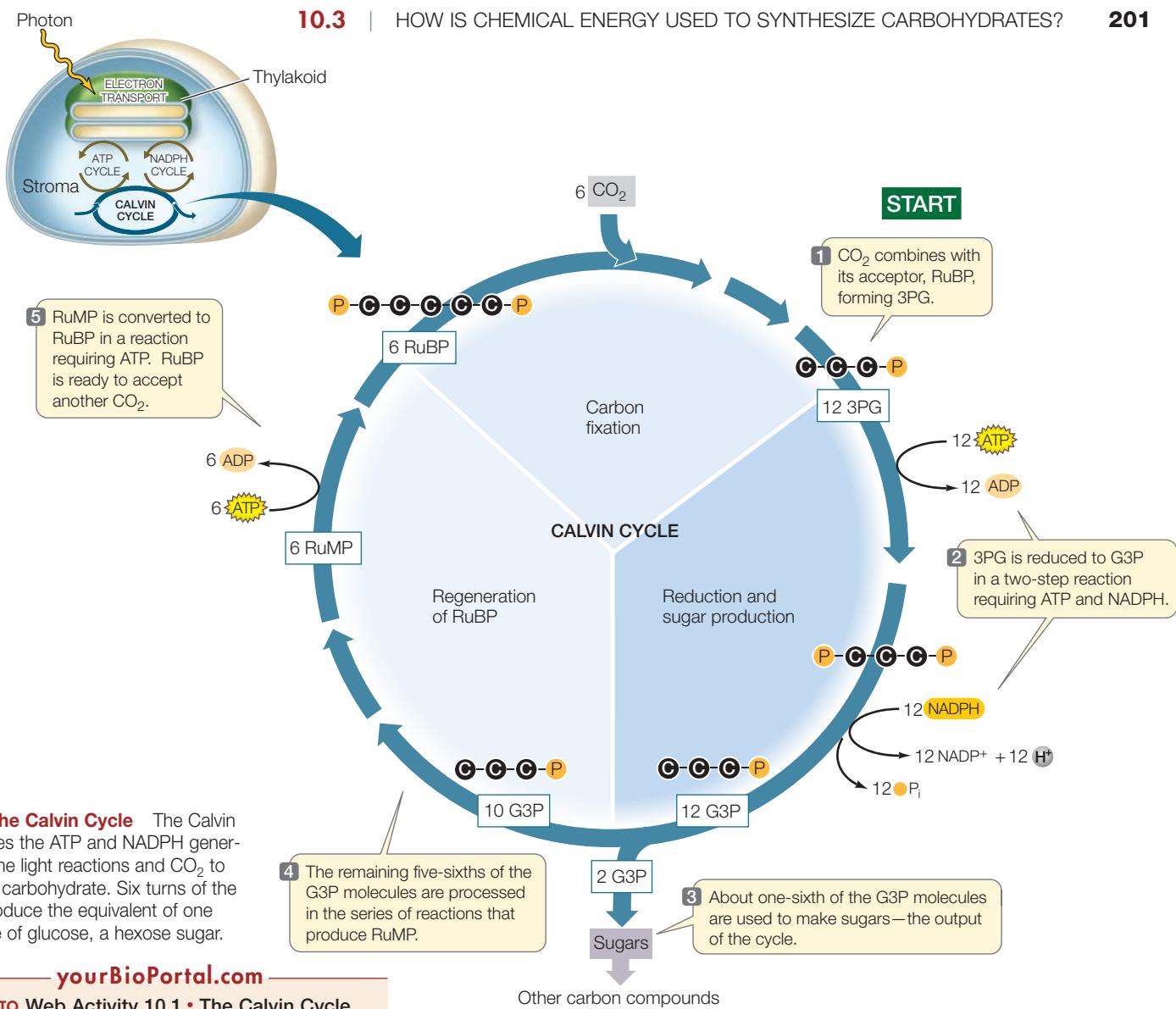
Glyceraldehyde 3-phosphate (G3P)

In a typical leaf, five-sixths of the G3P is recycled into RuBP. There are two fates for the remaining G3P, depending on the time of day and the needs of different parts of the plant:

- Some of it is exported out of the chloroplast to the cytosol, where it is converted to hexoses (glucose and fructose). These molecules may be used in glycolysis and mitochondrial respiration to power the activities of photosynthetic cells (see Chapter 9) or they may be converted into the disaccharide sucrose, which is transported out of the leaf to other organs in the plant. There it is hydrolyzed to its constituent monosaccharides, which can be used as sources of energy or as building blocks for other molecules.



10.14 RuBP Is the Carbon Dioxide Acceptor CO_2 is added to a 5-carbon compound, RuBP. The resulting 6-carbon compound immediately splits into two molecules of the sugar phosphate 3PG.



10.15 The Calvin Cycle The Calvin cycle uses the ATP and NADPH generated in the light reactions and CO_2 to produce carbohydrate. Six turns of the cycle produce the equivalent of one molecule of glucose, a hexose sugar.

- As the day wears on, glucose accumulates inside of the chloroplast, and these glucose units are linked to form the polysaccharide starch. This stored carbohydrate can then be drawn upon during the night so that the photosynthetic tissues can continue to export sucrose to the rest of the plant, even when photosynthesis is not taking place. In addition, starch is abundant in nonphotosynthetic organs such as roots, underground stems and seeds, where it provides a ready supply of glucose to fuel cellular activities, including plant growth.

The carbohydrates produced in photosynthesis are used by the plant to make other compounds. The carbon molecules are incorporated into amino acids, lipids, and the building blocks of nucleic acids—in fact all the organic molecules in the plant.

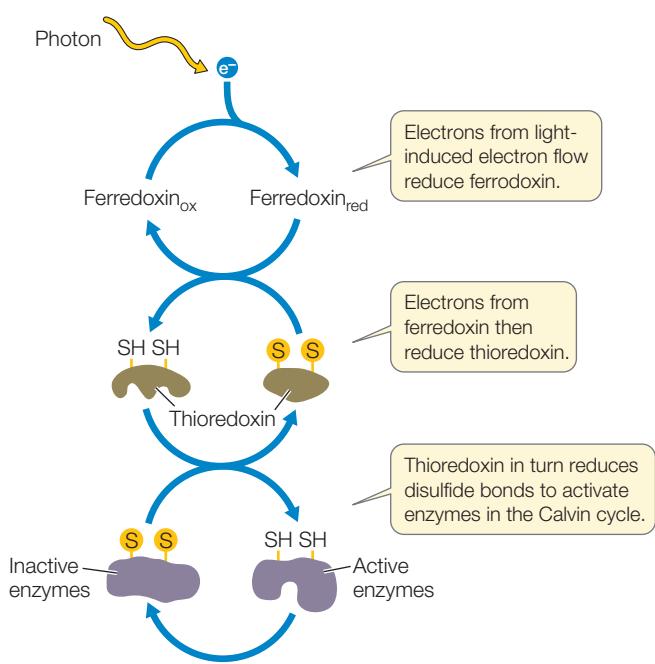
The products of the Calvin cycle are of crucial importance to the Earth's entire biosphere. For the majority of living organisms on Earth, the C—H covalent bonds generated by the cycle provide almost all of the energy for life. Photosynthetic organisms, which are also called **autotrophs** ("self-feeders"), release most of this energy by glycolysis and cellular respiration, and use it to support their own growth, development, and repro-

duction. But plants are also the source of energy for other organisms. Much plant matter ends up being consumed by **heterotrophs** ("other-feeders"), such as animals, which cannot photosynthesize. Heterotrophs depend on autotrophs for both raw materials and energy. Free energy is released from food by glycolysis and cellular respiration in heterotroph cells.

Light stimulates the Calvin cycle

As we have seen, the Calvin cycle uses NADPH and ATP, which are generated using energy from light. Two other processes connect the light reactions with this CO_2 fixation pathway. Both connections are indirect but significant:

- Light-induced pH changes in the stroma activate some Calvin cycle enzymes. Proton pumping from the stroma into the thylakoid lumen causes an increase in the pH of the stroma from 7 to 8 (a tenfold decrease in H^+ concentration). This favors the activation of rubisco.
- The light-induced electron transport reduces disulfide bonds in four of the Calvin cycle enzymes, thereby activat-



10.16 The Photochemical Reactions Stimulate the Calvin Cycle
By reducing (breaking) disulfide bridges, electrons from the light reactions activate enzymes in CO_2 fixation.

ing them (Figure 10.16). When ferredoxin is reduced in photosystem I (see Figure 10.10), it passes some electrons to a small, soluble protein called thioredoxin, and this protein passes electrons to four enzymes in the CO_2 fixation pathway. Reduction of the sulfurs in the disulfide bridges of these enzymes (see Figure 3.5) forms SH groups and breaks the bridges. The resulting changes in their three-dimensional shapes activate the enzymes and increase the rate at which the Calvin cycle operates.

10.3 RECAP

ATP and NADPH produced in the light reactions power the synthesis of carbohydrates by the Calvin cycle. This cycle fixes CO_2 , reduces it, and regenerates the acceptor, RuBP, for further fixation.

- Describe the experiments that led to the identification of RuBP as the initial CO_2 acceptor in photosynthesis.
See pp. 199–200 and Figure 10.13
- What are the three processes of the Calvin cycle?
See pp. 200–201 and Figure 10.15
- In what ways does light stimulate the Calvin cycle?
See pp. 201–202 and Figure 10.16

Although all green plants carry out the Calvin cycle, some plants have evolved variations on, or additional steps in, the light-independent reactions. These variations and additions have permitted plants to adapt to and thrive in certain environmental conditions. Let's look at these environmental limitations and the metabolic bypasses that have evolved to circumvent them.

10.4 How Do Plants Adapt to the Inefficiencies of Photosynthesis?

In addition to fixing CO_2 during photosynthesis, rubisco can react with O_2 . This reaction leads to a process called photorespiration, which lowers the overall rate of CO_2 fixation in some plants. After examining this problem, we'll look at some biochemical pathways and features of plant anatomy that compensate for the limitations of rubisco.

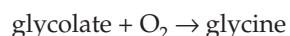
Rubisco catalyzes the reaction of RuBP with O_2 or CO_2

As its full name indicates, rubisco is an **oxygenase** as well as a **carboxylase**—it can add O_2 to the acceptor molecule RuBP instead of CO_2 . The affinity of rubisco for CO_2 is about ten times stronger than its affinity for O_2 . This means that inside a leaf with a normal exchange of air with the outside, CO_2 fixation is favored even though the concentration of CO_2 in the air is far less than that of O_2 . But if there is an even higher concentration of O_2 in the leaf, it acts as a competitive inhibitor, and RuBP reacts with O_2 rather than CO_2 . This reduces the overall amount of CO_2 that is converted into carbohydrates, and may play a role in limiting plant growth.

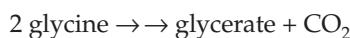
When O_2 is added to RuBP, one of the products is a 2-carbon compound, phosphoglycolate:



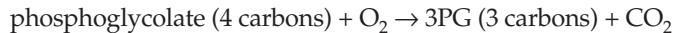
The 3PG formed by oxygenase activity enters the Calvin cycle but the phosphoglycolate does not. Plants have evolved a metabolic pathway that can partially recover the carbon in phosphoglycolate. The phosphoglycolate is hydrolyzed to glycolate, which diffuses into membrane-enclosed organelles called peroxisomes (Figure 10.17). There, a series of reactions converts it into the amino acid glycine:



The glycine then diffuses into a mitochondrion, where two glycine molecules are converted in a series of reactions into the amino acid serine, which goes back to the peroxisome and is converted into glycerate (a 3-carbon molecule) and CO_2 :

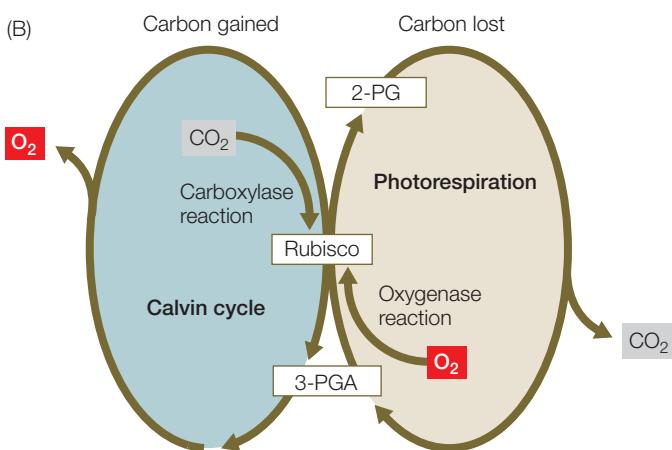
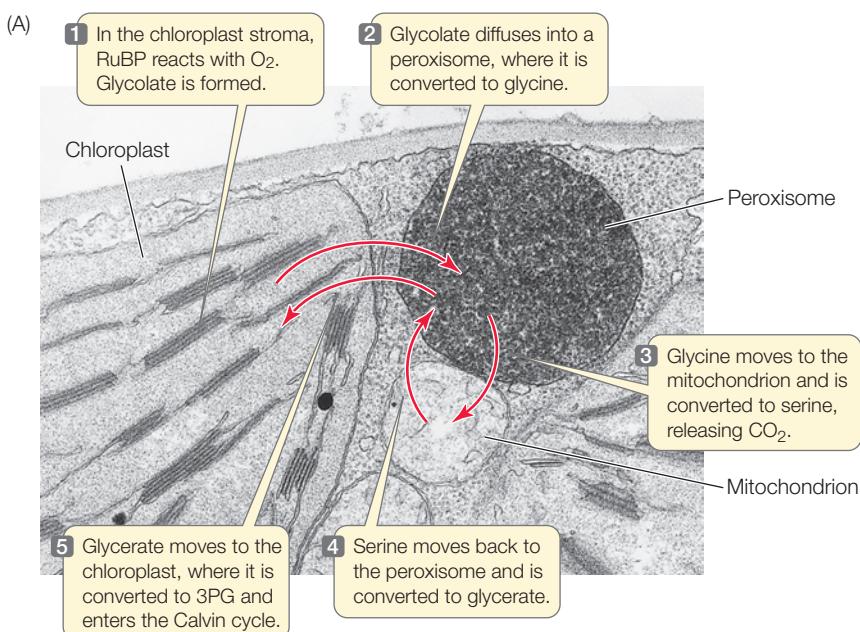


The glycerate moves into the chloroplast, where it is phosphorylated to make 3PG, which enters the Calvin cycle. So overall:



This pathway thus reclaims 75 percent of the carbons from phosphoglycolate for the Calvin cycle. In other words, the reaction of RuBP with O_2 instead of CO_2 reduces the net carbon fixed by the Calvin cycle by 25 percent. The pathway is called **photorespiration** because it consumes O_2 and releases CO_2 and because it occurs only in the light (due to the same enzyme activation processes that were mentioned above with regard to the Calvin cycle).

Why does rubisco act as an oxygenase as well as a carboxylase? Several factors are involved: active site affinities, concentrations of CO_2 and O_2 , and temperature.



10.17 Organelles of Photorespiration (A) The reactions of photorespiration take place in the chloroplasts, peroxisomes, and mitochondria. (B) Overall, photorespiration consumes O_2 and releases CO_2 .

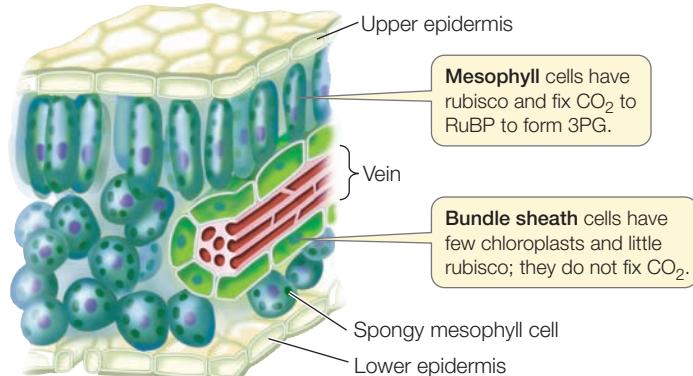
- As noted above, rubisco has a ten times higher affinity for CO_2 than for O_2 , and this favors CO_2 fixation.
- In the leaf, the relative concentrations of CO_2 and O_2 vary. If O_2 is relatively abundant, rubisco acts as an oxygenase and photorespiration ensues. If CO_2 predominates, rubisco fixes it for the Calvin cycle.
- Photorespiration is more likely at high temperatures. On a hot, dry day, small pores in the leaf surface called **stomata** close to prevent water from evaporating from the leaf (see Figure 10.1). But this also prevents gases from entering and leaving the leaf. The CO_2 concentration in the leaf falls as CO_2 is used up in photosynthetic reactions, and the O_2 concentration rises because of these same reactions. As the ratio of CO_2 to O_2 falls, the oxygenase activity of rubisco is favored, and photorespiration proceeds.

C₃ plants undergo photorespiration but C₄ plants do not

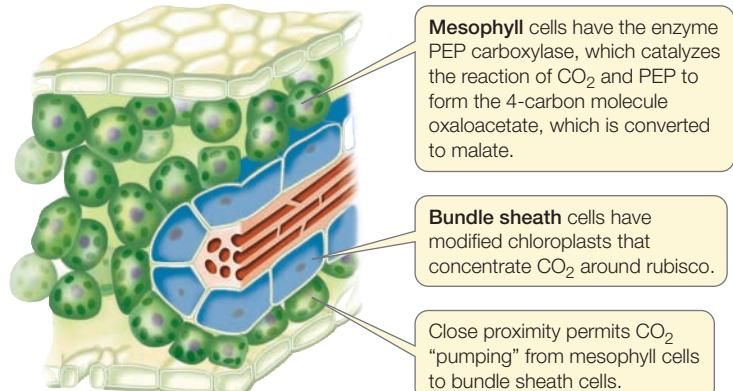
Plants differ in how they fix CO_2 , and can be distinguished as C₃ or C₄ plants, based on whether the first product of CO_2 fixation is a 3- or 4-carbon molecule. In **C₃ plants** such as roses, wheat, and rice, the first product is the 3-carbon molecule 3PG—as we have just described for the Calvin cycle. In these plants the cells of the mesophyll, which makes up the main body of the leaf, are full of chloroplasts containing rubisco (Figure 10.18A). On a hot day, these leaves close their stomata to conserve water, and as a result, rubisco acts as an oxygenase as well as a carboxylase, and photorespiration occurs.

C₄ plants, which include corn, sugarcane, and tropical grasses, make the 4-carbon molecule **oxaloacetate** as the first product of CO_2 fixation (Figure 10.18B). On a hot day, they partially close their stomata to conserve water, but their rate of photosynthesis does not fall, nor does photorespiration occur. What do they do differently?

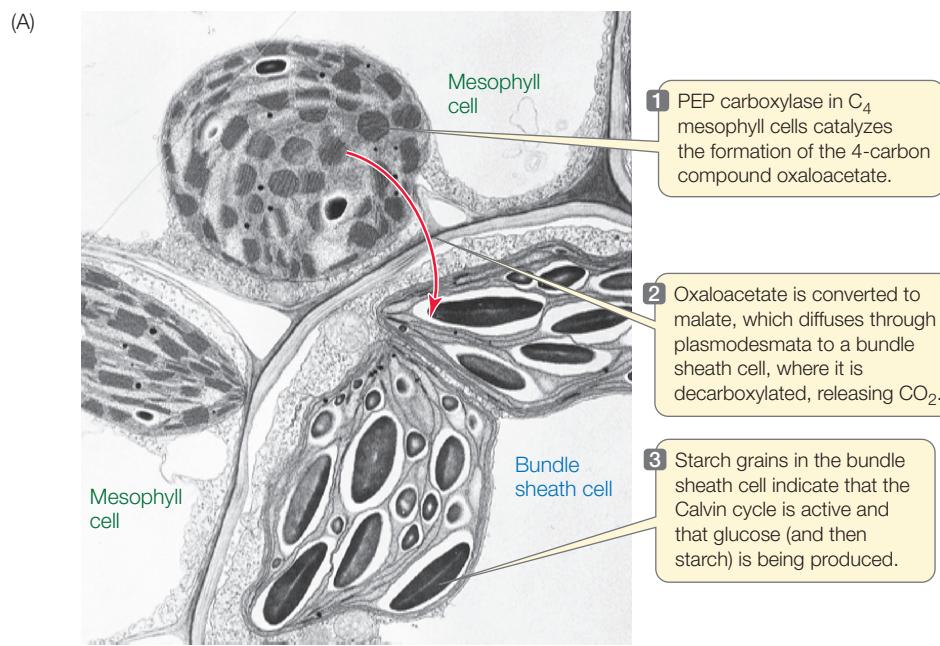
(A) Arrangement of cells in a C₃ leaf



(B) Arrangement of cells in a C₄ leaf



10.18 Leaf Anatomy of C₃ and C₄ Plants Carbon dioxide fixation occurs in different organelles and cells of the leaves in (A) C₃ plants and (B) C₄ plants. Cells that are tinted blue have rubisco.



10.19 The Anatomy and Biochemistry of C₄ Carbon Fixation (A) Carbon dioxide is fixed initially in the mesophyll cells, but enters the Calvin cycle in the bundle sheath cells. (B) The two cell types share an interconnected biochemical pathway for CO₂ assimilation.

surface of the leaf. It fixes CO₂ to a 3-carbon acceptor compound, **phosphoenolpyruvate (PEP)**, to produce the 4-carbon fixation product, oxaloacetate. PEP carboxylase has two advantages over rubisco:

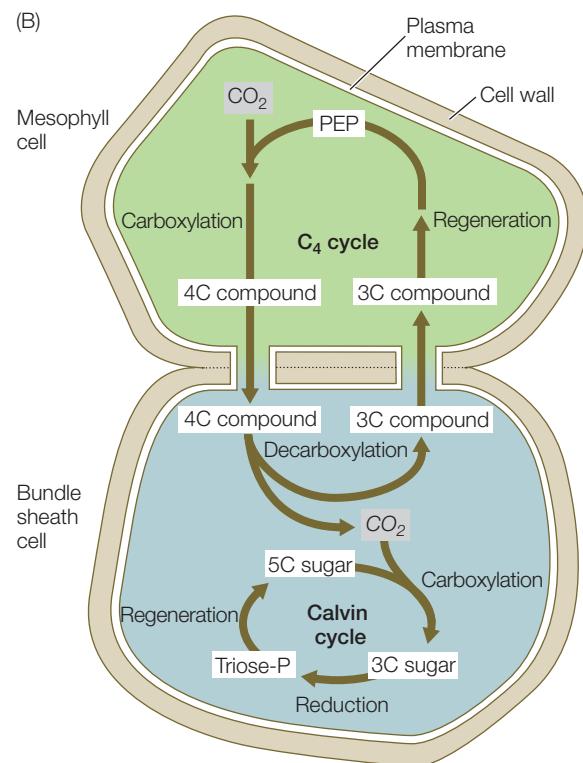
- It does not have oxygenase activity.
- It fixes CO₂ even at very low CO₂ levels.

So even on a hot day when the stomata are partially closed and the ratio of O₂ to CO₂ rises, PEP carboxylase just keeps on fixing CO₂.

Oxaloacetate is converted to malate, which diffuses out of the mesophyll cells and into the **bundle sheath cells** (see Figure 10.18B), located in the interior of the leaf. (Some C₄ plants convert the oxaloacetate to aspartate instead of malate, but we will only discuss the malate pathway here.) The bundle sheath cells contain modified chloroplasts that are designed to concentrate CO₂ around the rubisco. There, the 4-carbon malate loses one carbon (is decarboxylated), forming CO₂ and pyruvate. The latter moves back to the mesophyll cells where the 3-carbon acceptor compound, PEP, is regenerated at the expense of ATP. Thus the role of PEP is to bind CO₂ from the air in the leaf so that it can be transferred to the bundle sheath cells, where it is delivered to rubisco. This process essentially “pumps up” the CO₂ concentration around rubisco, so that it acts as a carboxylase and begins the Calvin cycle.

C₃ plants have an advantage over C₄ plants in that they don’t expend extra ATP to “pump up” the concentration of CO₂ near rubisco. But this advantage begins to be outweighed under conditions that favor photorespiration, such as warmer seasons and climates. Under these conditions C₄ plants have the advantage. For example, Kentucky bluegrass is a C₃ plant that thrives on lawns in April and May. But in the heat of summer it does not do as well and Bermuda grass, a C₄ plant, takes over the lawn. The same is true on a global scale for crops: C₃ plants such as soybean, wheat, and barley have been adapted for human food production in temperate climates, while C₄ plants such as corn and sugarcane originated and are grown in the tropics.

THE EVOLUTION OF CO₂ FIXATION PATHWAYS C₃ plants are certainly more ancient than C₄ plants. While C₃ photosynthesis appears to have begun about 3.5 billion years ago, C₄ plants appeared about 12 million years ago. A possible factor in the emergence of the C₄ pathway is the decline in atmospheric CO₂. When dinosaurs dominated Earth 100 million years ago, the concentration of CO₂ in the atmosphere was four times what it is now. As CO₂ levels declined thereafter, the more efficient C₄



C₄ plants perform the normal Calvin cycle, but they have an additional early reaction that fixes CO₂ without losing carbon to photorespiration. Because this initial CO₂ fixation step can function even at low levels of CO₂ and high temperatures, C₄ plants very effectively optimize photosynthesis under conditions that inhibit it in C₃ plants. C₄ plants have *two separate enzymes for CO₂ fixation, located in different parts of the leaf* (Figure 10.19; see also Figure 10.18B). The first enzyme, called **PEP carboxylase**, is present in the cytosols of mesophyll cells near the

TABLE 10.1
Comparison of Photosynthesis in C₃, C₄, and CAM Plants

	C ₃ PLANTS	C ₄ PLANTS	CAM PLANTS
Calvin cycle used?	Yes	Yes	Yes
Primary CO ₂ acceptor	RuBP	PEP	PEP
CO ₂ -fixing enzyme	Rubisco	PEP carboxylase	PEP carboxylase
First product of CO ₂ fixation	3PG (3-carbon)	Oxaloacetate (4-carbon)	Oxaloacetate (4-carbon)
Affinity of carboxylase for CO ₂	Moderate	High	High
Photosynthetic cells of leaf	Mesophyll	Mesophyll and bundle sheath	Mesophyll with large vacuoles
Photorespiration	Extensive	Minimal	Minimal

plants would have gained an advantage over their C₃ counterparts.

As we described in the opening essay of this chapter, CO₂ levels have been increasing over the past 200 years. Currently, the level of CO₂ is not enough for maximal CO₂ fixation by rubisco, so photorespiration occurs, reducing the growth rates of C₃ plants. Under hot conditions, C₄ plants are favored. But if CO₂ levels in the atmosphere continue to rise, the reverse will occur and C₃ plants will have a comparative advantage. The overall growth rates of crops such as rice and wheat should increase. This may or may not translate into more food, given that other effects of the human-spurred CO₂ increase (such as global warming) will also alter Earth's ecosystems.

CAM plants also use PEP carboxylase

Other plants besides the C₄ plants use PEP carboxylase to fix and accumulate CO₂. They include some water-storing plants (called succulents) of the family Crassulaceae, many cacti, pineapples, and several other kinds of flowering plants. The CO₂ metabolism of these plants is called **crassulacean acid metabolism**, or **CAM**, after the family of succulents in which it was discovered. CAM is much like the metabolism of C₄ plants in that CO₂ is initially fixed into a 4-carbon compound. But in CAM plants the initial CO₂ fixation and the Calvin cycle are separated in time rather than space.

- At night, when it is cooler and water loss is minimized, the stomata open. CO₂ is fixed in mesophyll cells to form the 4-carbon compound oxaloacetate, which is converted into malate and stored in the vacuole.
- During the day, when the stomata close to reduce water loss, the accumulated malate is shipped to the chloroplasts, where its decarboxylation supplies the CO₂ for the Calvin cycle and the light reactions supply the necessary ATP and NADPH.

CAM benefits the plant by allowing it to close its stomata during the day. As you will learn in Chapter 35, plants lose most of the water that they take up in their roots by evaporation through the leaves (transpiration). In dry climates, closing stomata is a key to water conservation and survival.

Table 10.1 compares photosynthesis in C₃, C₄, and CAM plants.

10.4 RECAP

Rubisco catalyzes the carboxylation of RuBP to form two 3PG, and the oxygenation of RuBP to form one 3PG and one phosphoglycolate. The diversion of rubisco to its oxygenase function decreases net CO₂ fixation. C₄ photosynthesis and CAM allow plants to adapt to environmental conditions that result in a limited availability of CO₂ inside the leaf.

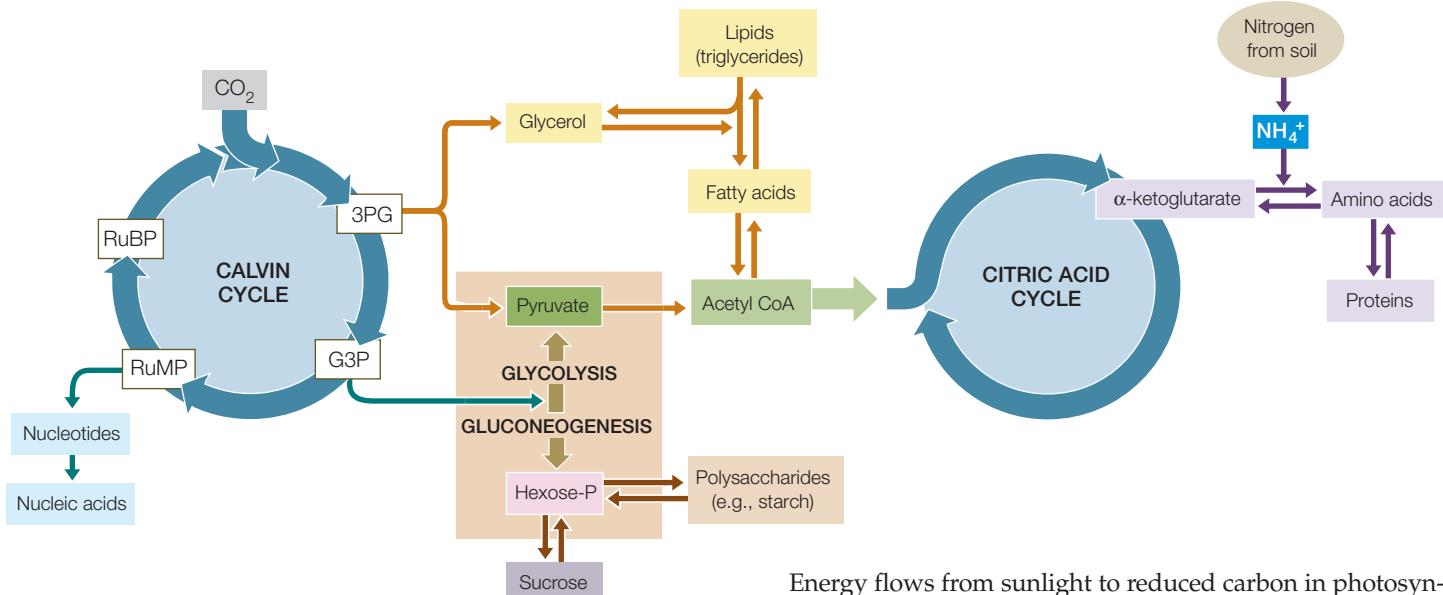
- Explain how photorespiration recovers some of the carbon that is channeled away from the Calvin cycle. **See pp. 202–203 and Figure 10.17**
- What do C₄ plants do to keep the concentration of CO₂ around rubisco high, and why? **See pp. 203–204 and Figure 10.19**
- What is the pathway for CO₂ fixation in CAM plants? **See p. 205**

Now that we understand how photosynthesis produces carbohydrates, let's see how the pathways of photosynthesis are connected to other metabolic pathways.

10.5 How Does Photosynthesis Interact with Other Pathways?

Green plants are autotrophs and can synthesize all the molecules they need from simple starting materials: CO₂, H₂O, phosphate, sulfate, ammonium ions (NH₄⁺), and small quantities of other mineral nutrients. The NH₄⁺ is needed to synthesize amino acids and nucleotides, and it comes from either the conversion of nitrogen-containing molecules in soil water or the conversion of N₂ gas from the atmosphere by bacteria, as we will see in Chapter 36.

Plants use the carbohydrates generated in photosynthesis to provide energy for processes such as active transport and anabolism. Both cellular respiration and fermentation can occur in plants, although the former is far more common. Unlike photosynthesis, plant cellular respiration takes place both in the light and in the dark.



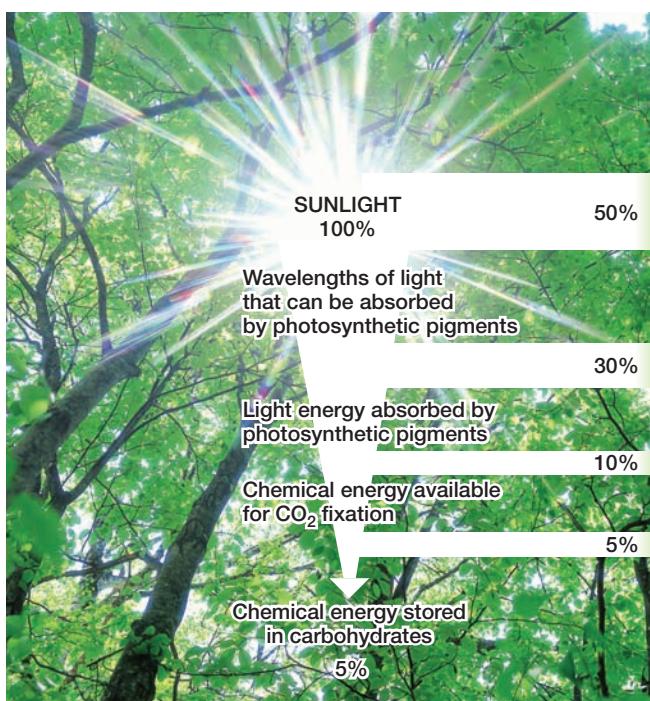
10.20 Metabolic Interactions in a Plant Cell The products of the Calvin cycle are used in the reactions of cellular respiration (glycolysis and the citric acid cycle).

Photosynthesis and respiration are closely linked through the Calvin cycle (**Figure 10.20**). The partitioning of G3P is particularly important:

- Some G3P from the Calvin cycle takes part in the glycolysis pathway and is converted into pyruvate in the cytosol. This pyruvate can be used in cellular respiration for energy, or its carbon skeletons can be used in anabolic reactions to make lipids, proteins, and other carbohydrates (see Figure 9.14).
- Some G3P can enter a pathway that is the reverse of glycolysis (gluconeogenesis; see Section 9.5). In this case, hexose-phosphates and then sucrose are formed and transported to the nonphotosynthetic tissues of the plant (such as the root).

Energy flows from sunlight to reduced carbon in photosynthesis, then to ATP in respiration. Energy can also be stored in the bonds of macromolecules such as polysaccharides, lipids, and proteins. For a plant to grow, energy storage (as body structures) must exceed energy release; that is, overall carbon fixation by photosynthesis must exceed respiration. This principle is the basis of the ecological food chain, as we will see in later chapters.

Photosynthesis provides most of the energy that we need for life. Given the uncertainties about the future of photosynthesis (due to changes in CO₂ levels and climate change), it would be wise to seek ways to improve photosynthetic efficiency. **Figure 10.21** shows the various ways in which solar energy is utilized by plants or lost. In essence, only about 5 percent of the sunlight that reaches Earth is converted into plant growth. The inefficiencies of photosynthesis involve basic chemistry and physics (some light energy is not absorbed by photosynthetic pigments) as well as biology (plant anatomy and leaf exposure, photorespiration, and inefficiencies in metabolic pathways). While it is hard to change chemistry and physics, biologists might be able to use their knowledge of plants to improve on the basic biology of photosynthesis. This could result in a more efficient use of resources and better food production.



10.5 RECAP

The products of photosynthesis are utilized in glycolysis and the citric acid cycle, as well as in the synthesis of lipids, proteins, and other large molecules.

- How do common intermediates link the pathways of glycolysis, the citric acid cycle, and photosynthesis? See p. 206 and Figure 10.20

10.21 Energy Losses During Photosynthesis As we face an increasingly uncertain future for photosynthesis on Earth, understanding its inefficiencies becomes increasingly important. Photosynthetic pathways preserve at most about 5 percent of the sun's energy input as chemical energy in carbohydrates.

CHAPTER SUMMARY

10.1 What Is Photosynthesis?

- In the process of **photosynthesis**, plants and other organisms take in CO₂, water, and light energy, producing O₂ and carbohydrates. **SEE ANIMATED TUTORIAL 10.1**
- The **light reactions** of photosynthesis convert light energy into chemical energy. They produce ATP and reduce NADP⁺ to NADPH. **Review Figure 10.3**
- The **light-independent reactions** do not use light directly but instead use ATP and NADPH to reduce CO₂, forming carbohydrates.

10.2 How Does Photosynthesis Convert Light Energy into Chemical Energy?

- Light is a form of electromagnetic radiation. It is emitted in particle-like packets called **photons** but has wavelike properties.
- Molecules that absorb light in the visible spectrum are called **pigments**. Photosynthetic organisms have several pigments, most notably **chlorophylls**, but also accessory pigments such as carotenoids and phycobilins.
- Absorption of a photon puts a pigment molecule in an excited state that has more energy than its ground state. **Review Figure 10.5**
- Each compound has a characteristic **absorption spectrum**. An **action spectrum** reflects the biological activity of a photosynthetic organism for a given wavelength of light. **Review Figure 10.6**
- The pigments in photosynthetic organisms are arranged into **antenna systems** that absorb energy from light and funnel this energy to a pair of chlorophyll *a* molecules in the reaction center of the **photosystem**. Chlorophyll can act as a reducing agent, transferring excited electrons to other molecules. **Review Figure 10.8**
- Noncyclic electron transport** uses photosystems I and II to produce ATP, NADPH and O₂. **Cyclic electron transport** uses only photosystem I and produces only ATP. **Review Figures 10.10 and 10.11**
- Chemiosmosis is the mechanism of ATP production in **photophosphorylation**. **Review Figure 10.12, ANIMATED TUTORIAL 10.2**

10.3 How Is Chemical Energy Used to Synthesize Carbohydrates?

- The **Calvin cycle** makes carbohydrates from CO₂. The cycle consists of three processes: fixation of CO₂, reduction and carbohydrate production, and regeneration of RuBP. **SEE ANIMATED TUTORIAL 10.3**
- RuBP** is the initial CO₂ acceptor, and **3PG** is the first stable product of CO₂ fixation. The enzyme **rubisco** catalyzes the reaction of CO₂ and RuBP to form 3PG. **Review Figure 10.14, WEB ACTIVITY 10.1**
- ATP and NADPH formed by the light reactions are used in the reduction of 3PG to form **G3P**. **Review Figure 10.15**
- Light stimulates enzymes in the Calvin cycle, further integrating the light-dependent and light-independent pathways.

10.4 How Do Plants Adapt to the Inefficiencies of Photosynthesis?

- Rubisco can catalyze a reaction between O₂ and RuBP in addition to the reaction between CO₂ and RuBP. At high temperatures and low CO₂ concentrations, the **oxygenase** function of rubisco is favored over its **carboxylase** function.
- When rubisco functions as an oxygenase, the result is **photo-respiration**, which significantly reduces the efficiency of photosynthesis.
- In **C₄ plants**, CO₂ reacts with PEP to form a 4-carbon intermediate in mesophyll cells. The 4-carbon product releases its CO₂ to rubisco in the **bundle sheath** cells in the interior of the leaf. **Review Figure 10.18, WEB ACTIVITY 10.2**
- CAM** plants operate much like C₄ plants, but their initial CO₂ fixation by PEP carboxylase is temporally separated from the Calvin cycle, rather than spatially separated as in C₄ plants.

10.5 How Does Photosynthesis Interact with Other Pathways?

- Photosynthesis and cellular respiration are linked through the **Calvin cycle**, the **citric acid cycle**, and **glycolysis**. **Review Figure 10.20**
- To survive, a plant must photosynthesize more than it respires.
- Photosynthesis utilizes only a small portion of the energy of sunlight. **Review Figure 10.21**

SELF-QUIZ

- In noncyclic photosynthetic electron transport, water is used to
 - excite chlorophyll.
 - hydrolyze ATP.
 - reduce P_i.
 - oxidize NADPH.
 - reduce chlorophyll.
- Which statement about light is true?
 - An absorption spectrum is a plot of biological effectiveness versus wavelength.
 - An absorption spectrum may be a good means of identifying a pigment.
 - Light need not be absorbed to produce a biological effect.
 - A given kind of molecule can occupy any energy level.
 - A pigment loses energy as it absorbs a photon.
- Which statement about chlorophylls is *not* true?
 - Chlorophylls absorb light near both ends of the visible spectrum.
 - Chlorophylls can accept energy from other pigments, such as carotenoids.
 - Excited chlorophyll can either reduce another substance or release light energy.
 - Excited chlorophyll cannot be an oxidizing agent.
 - Chlorophylls contain magnesium.

4. In cyclic electron transport,
 - a. oxygen gas is released.
 - b. ATP is formed.
 - c. water donates electrons and protons.
 - d. NADPH forms.
 - e. CO₂ reacts with RuBP.
5. Which of the following does *not* happen in noncyclic electron transport?
 - a. Oxygen gas is released.
 - b. ATP forms.
 - c. Water donates electrons and protons.
 - d. NADPH forms.
 - e. CO₂ reacts with RuBP.
6. In chloroplasts,
 - a. light leads to the flow of protons out of the thylakoids.
 - b. ATP is formed when protons flow into the thylakoid lumen.
 - c. light causes the thylakoid lumen to become less acidic than the stroma.
 - d. protons return passively to the stroma through protein channels.
 - e. proton pumping requires ATP.
7. Which statement about the Calvin cycle is *not* true?
 - a. CO₂ reacts with RuBP to form 3PG.
 - b. RuBP forms by the metabolism of 3PG.
8. In C₄ photosynthesis,
 - a. 3PG is the first product of CO₂ fixation.
 - b. rubisco catalyzes the first step in the pathway.
 - c. 4-carbon acids are formed by PEP carboxylase in bundle sheath cells.
 - d. photosynthesis continues at lower CO₂ levels than in C₃ plants.
 - e. CO₂ released from RuBP is transferred to PEP.
9. Photosynthesis in green plants occurs only during the day. Respiration in plants occurs
 - a. only at night.
 - b. only when there is enough ATP.
 - c. only during the day.
 - d. all the time.
 - e. in the chloroplast after photosynthesis.
10. Photorespiration
 - a. takes place only in C₄ plants.
 - b. includes reactions carried out in peroxisomes.
 - c. increases the yield of photosynthesis.
 - d. is catalyzed by PEP carboxylase.
 - e. is independent of light intensity.

FOR DISCUSSION

1. Both photosynthetic electron transport and the Calvin cycle stop in the dark. Which specific reaction stops first? Which stops next? Continue answering the question “Which stops next?” until you have explained why both pathways have stopped.
2. In what principal ways are the reactions of electron transport in photosynthesis similar to the reactions of oxidative phosphorylation discussed in Section 9.3?
3. Differentiate between cyclic and noncyclic electron transport in terms of (1) the products and (2) the source of electrons for the reduction of oxidized chlorophyll.
4. If water labeled with ¹⁸O is added to a suspension of photosynthesizing chloroplasts, which of the following compounds will first become labeled with ¹⁸O: ATP, NADPH, O₂, or 3PG? If water labeled with ³H is added, which of the same compounds will first become radioactive? Which will be first if CO₂ labeled with ¹⁴C is added?
5. The Viking lander was sent to Mars in 1976 to detect signs of life. Explain the rationale behind the following experiments this unmanned probe performed:
 - a. A scoop of dirt was inserted into a container and ¹⁴CO₂ was added. After a while during the Martian day, the ¹⁴CO₂ was removed and the dirt was heated to a high temperature. Scientists monitoring the experiment back on Earth looked for the release of ¹⁴CO₂ as a sign of life.
 - b. The same experiment was performed, except that the dirt was heated to a high temperature for 30 minutes and then allowed to cool to Martian temperature right after scooping, and before the ¹⁴CO₂ was added. If living things were present, then ¹⁴CO₂ would be released in experiment (a), but not this one.

ADDITIONAL INVESTIGATION

Calvin's experiment (see Figure 10.13) laid the foundations for a full description of the pathway of CO₂ fixation. Given the interrelationships between metabolic pathways in plants, how

would you do an experiment to follow the pathway of fixed carbon from photosynthesis to proteins?

WORKING WITH DATA (GO TO yourBioPortal.com)

Water is the Source of the Oxygen Produced by Photosynthesis

The proposal that the source of O₂ in photosynthesis was H₂O rather than CO₂ was first made in 1932. But it took the invention of isotope tracing a decade later to prove this. In this exercise, you will examine the methods that Ruben and Kamen used (Figure 10.2) to identify the isotopes of oxygen and the data they obtained.

Tracing the Pathway of CO₂ Studies of radioactive isotopes were intensified during World War II as an offshoot of the development of nuclear weapons. This led Calvin and his colleagues to perform the experiments designed to trace the path of carbon in photosynthesis (Figure 10.13). In this hands-on exercise, you will examine their data and see the reasoning that led to the CO₂ fixation pathway.

11

The Cell Cycle and Cell Division

An enemy of the cell reproduction cycle

Ruth felt healthy and was surprised when she was called back to her physician's office a week after her annual checkup. "Your lab report indicates you have early cervical cancer," said the doctor. "I ordered a follow-up test, and it came back positive—at some point, you were infected with HPV."

Ruth felt numb as soon as she heard the word "cancer." Her mother had died of breast cancer in the previous year. The doctor's statement about HPV (human papillomavirus) did not register in her consciousness. Sensing Ruth's discomfort, the doctor quickly reassured her that the cancer was caught at an early stage, and that a simple surgical procedure would remove it. Two weeks later, the cancer was removed and Ruth remains cancer-free. She was fortunate that her annual medical exam included a Papanicolaou (Pap) test, in which the cells lining the cervix

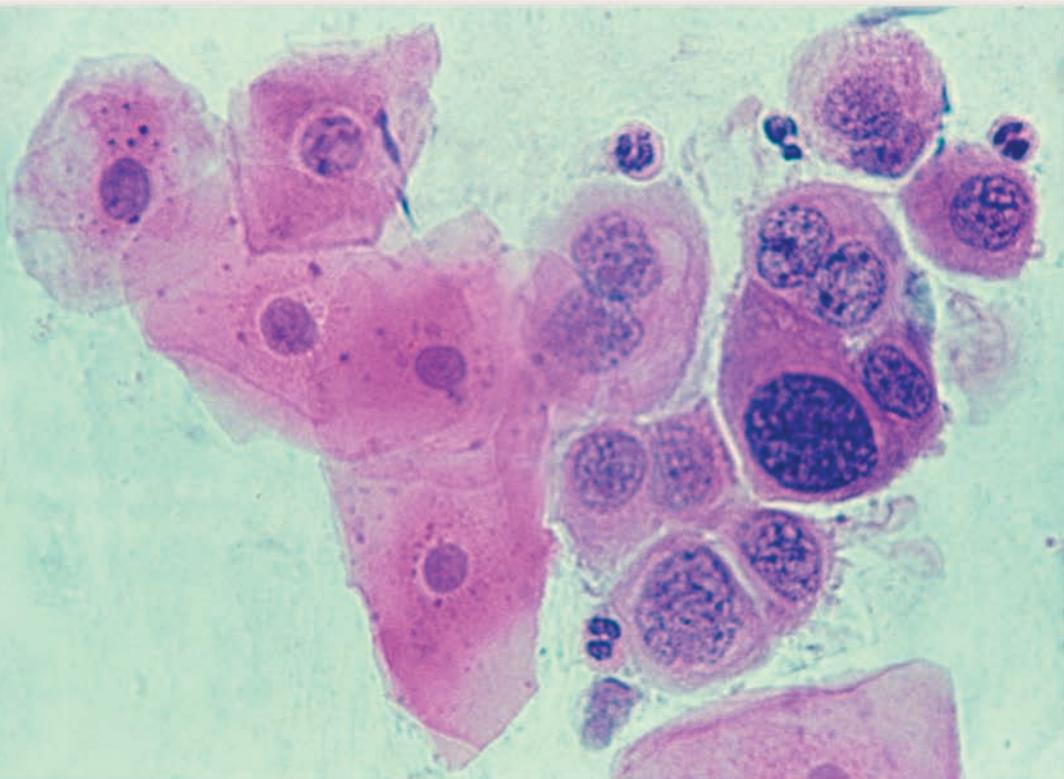
are examined for abnormalities. Since they were begun almost 50 years ago in Europe, Pap tests have resulted in the early detection and removal of millions of early cervical cancers, and the death rate from this potentially lethal disease has plummeted.

Only recently was HPV found to be the cause of most cervical cancers. The German physician Harald Zur-Hansen was awarded the Nobel Prize in 2008 for this discovery and it has led to a vaccine to prevent future infections. There are many different types of HPV and many of the ones that infect humans cause warts, which are small, rough growths on the skin. The types of HPV that infect tissues at the cervix get there by sexual transmission, and this is a common infection in Western societies.

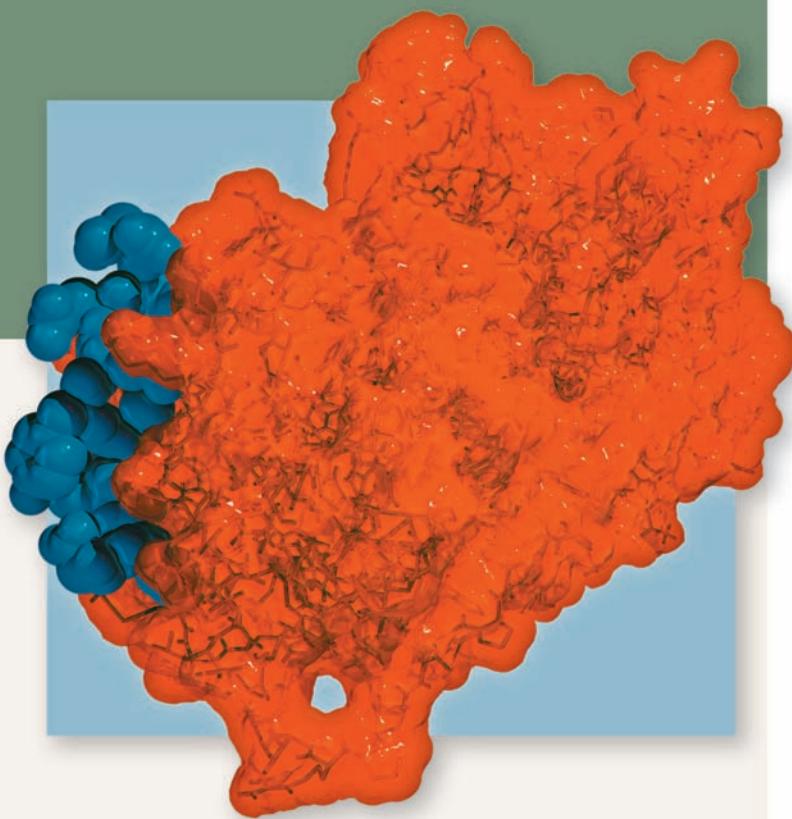
When HPV arrives at the tissues lining the cervix, it has one of two fates. Most of the time, it gets into the cells,

turning them into HPV factories, producing a lot of HPV in the mucus outside the uterus. These viruses can infect another person during a sexual encounter. In some cases, however, the virus follows a different and—for the host cells—more sinister path. It infects the cervical cells and causes them to make a viral protein called E7, a protein that can deregulate human cell reproduction.

Cell reproduction in healthy humans is tightly controlled, and one of the strongest regulators that prevent a cell from dividing is the



Abnormal Cells In this Pap test, cervical cancer cells at right differ from the normal cells at left. The cancer cells have larger nuclei.



E7, RB, and Cell Reproduction The E7 protein (blue) from human papillomavirus binds to the RB protein (red) to inhibit RB's ability to block cell division. This results in cancer.

retinoblastoma protein (RB), which you will encounter later in the chapter. One of the viral gene products is the protein E7, which has a three-dimensional shape that just fits into the protein-binding site of RB, thereby inactivating it. With no active RB to put the brakes on, cell division proceeds. As you know, uncontrolled cell reproduction is a hallmark of cancer—and so cervical cancer begins.

Understanding the cell division cycle and its control is clearly an important subject for understanding cancer. But cell division is not just important in medicine. It underlies the growth, development and reproduction of all organisms.

IN THIS CHAPTER we will see how cells give rise to more cells. We first describe how prokaryotic cells divide to produce new, single-celled organisms. Then we turn to the two types of nuclear division in eukaryotes—mitosis and meiosis—and relate them to asexual and sexual reproduction. Cell reproduction is linked to cell death, so we then consider the process of programmed cell death, also known as apoptosis. Finally, we relate these processes to the loss of cell reproduction control in cancer cells.

CHAPTER OUTLINE

- 11.1 How Do Prokaryotic and Eukaryotic Cells Divide?
- 11.2 How Is Eukaryotic Cell Division Controlled?
- 11.3 What Happens during Mitosis?
- 11.4 What Role Does Cell Division Play in a Sexual Life Cycle?
- 11.5 What Happens during Meiosis?
- 11.6 In a Living Organism, How Do Cells Die?
- 11.7 How Does Unregulated Cell Division Lead to Cancer?

11.1 How Do Prokaryotic and Eukaryotic Cells Divide?

The life cycle of an organism, from birth to death, is intimately linked to cell division. Cell division plays important roles in the growth and repair of tissues in multicellular organisms, as well as in the reproduction of all organisms (Figure 11.1).

In order for any cell to divide, four events must occur:

- There must be a reproductive signal. This signal initiates cell division and may originate from either inside or outside the cell.
- **Replication** of DNA (the genetic material) must occur so that each of the two new cells will have identical genes and complete cell functions.
- The cell must distribute the replicated DNA to each of the two new cells. This process is called **segregation**.
- In addition to synthesizing needed enzymes and organelles, new material must be added to the plasma membrane (and the cell wall, in organisms that have one), in order to separate the two new cells by a process called **cytokinesis**.

These four events proceed somewhat differently in prokaryotes and eukaryotes.

Prokaryotes divide by binary fission

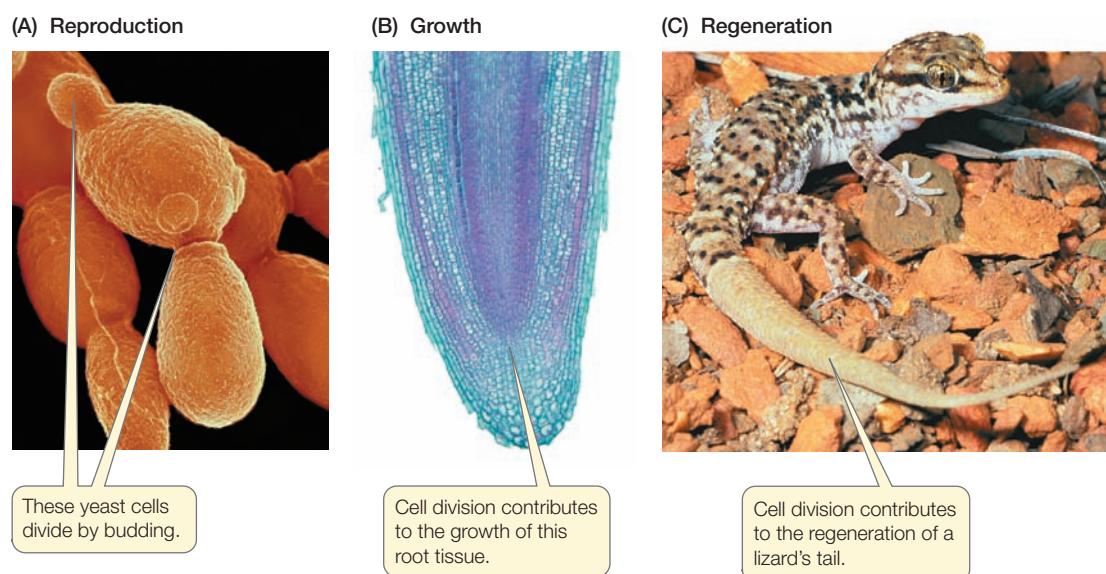
In prokaryotes, cell division results in the reproduction of the entire single-celled organism. The cell grows in size, replicates its DNA, and then separates the cytoplasm and DNA into two new cells by a process called **binary fission**.

REPRODUCTIVE SIGNALS The reproductive rates of many prokaryotes respond to conditions in the environment. The bacterium *Escherichia coli*, a species commonly used in genetic studies, is a “cell division machine”; if abundant sources of carbohydrates and mineral nutrients are available, it can divide as often as every 20 minutes. Another bacterium, *Bacillus subtilis*, does not just slow its growth when nutrients are low but stops dividing and then resumes dividing when conditions improve. Clearly, external factors such as environmental conditions and nutrient concentrations are signals for the initiation of cell division in prokaryotes.

REPLICATION OF DNA As we saw in Section 5.3, a **chromosome** can be defined in molecular terms as a DNA molecule containing genetic information. When a cell divides, all of its chromo-

11.1 Important Consequences of Cell Division

Division Cell division is the basis for (A) reproduction, (B) growth, and (C) repair and regeneration of tissues.



somes must be replicated, and one copy of each chromosome must find its way into one of the two new cells.

Most prokaryotes have only one chromosome—a single long DNA molecule with proteins bound to it. In *E. coli*, the ends of the DNA molecule are joined to create a circular chromosome. Circular chromosomes are characteristic of most prokaryotes as well as some viruses, and are also found in the chloroplasts and mitochondria of eukaryotic cells.

If the *E. coli* DNA were spread out into an actual circle, it would be about 500 μm in diameter. The bacterium itself is only about 2 μm long and 1 μm in diameter. Thus if the bacterial DNA were fully extended, it would form a circle over 200 times larger than the cell! To fit into the cell, bacterial DNA must be compacted. The DNA folds in on itself, and positively charged (basic) proteins bound to the negatively charged (acidic) DNA contribute to this folding.

Two regions of the prokaryotic chromosome play functional roles in cell reproduction:

- *ori*: the site where replication of the circular chromosome starts (the *origin of replication*)
- *ter*: the site where replication ends (the *terminus of replication*)

Chromosome replication takes place as the DNA is threaded through a “replication complex” of proteins near the center of the cell. (These proteins include the enzyme DNA polymerase, whose important role in replication is discussed further in Section 13.3.) Replication begins at the *ori* site and moves toward the *ter* site. While the DNA replicates, anabolic metabolism is active and the cell grows. When replication is complete, the two daughter DNA molecules separate and segregate from one another at opposite ends of the cell. In rapidly dividing prokaryotes, DNA replication occupies the entire time between cell divisions.

SEGREGATION OF DNAs Replication begins near the center of the cell, and as it proceeds, the *ori* regions move toward opposite ends of the cell (Figure 11.2A). DNA sequences adjacent to the

ori region bind proteins that are essential for this segregation. This is an active process, since the binding proteins hydrolyze ATP. The prokaryotic cytoskeleton (see Section 5.2) may be involved in DNA segregation, either actively moving the DNA along, or passively acting as a “railroad track” along which DNA moves.

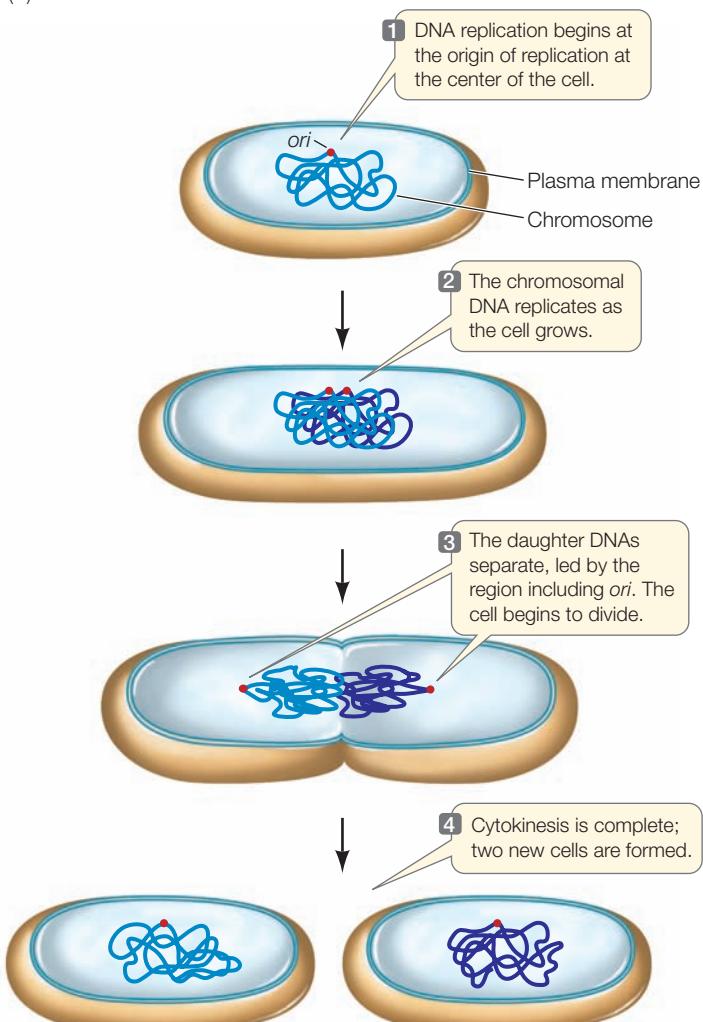
CYTOKINESIS The actual division of a single cell and its contents into two cells is called cytokinesis and begins immediately after chromosome replication is finished in rapidly growing cells. The first event of cytokinesis is a pinching in of the plasma membrane to form a ring of fibers similar to a purse string. The major component of these fibers is a protein that is related to eukaryotic tubulin (which makes up microtubules). As the membrane pinches in, new cell wall materials are deposited, which finally separate the two cells (Figure 11.2B).

Eukaryotic cells divide by mitosis or meiosis followed by cytokinesis

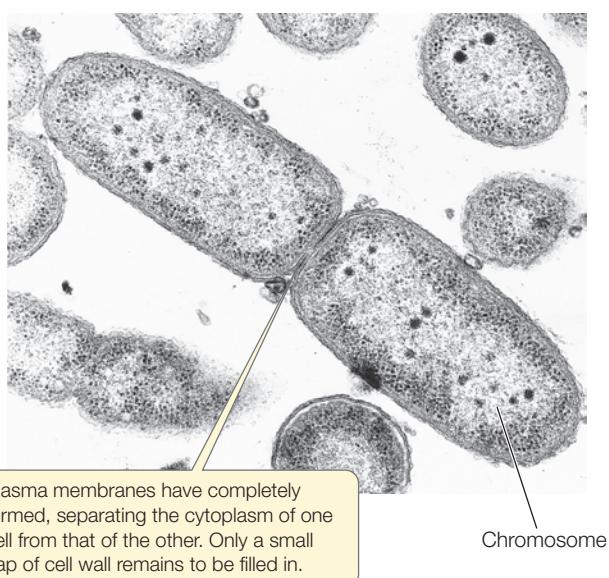
As in prokaryotes, cell reproduction in eukaryotes entails reproductive signals, DNA replication, segregation, and cytokinesis. The details, however, are quite different:

- **Signal.** Unlike prokaryotes, eukaryotic cells do not constantly divide whenever environmental conditions are adequate. In fact, most eukaryotic cells that are part of a multicellular organism and have become specialized seldom divide. In a eukaryotic organism, the signals for cell division are related not to the environment of a single cell, but to the needs of the entire organism.
- **Replication.** While most prokaryotes have a single main chromosome, eukaryotes usually have many (humans have 46). Consequently the processes of replication and segregation are more intricate in eukaryotes than in prokaryotes. In eukaryotes, DNA replication is usually limited to a portion of the period between cell divisions.

(A)



(B)



11.2 Prokaryotic Cell Division (A) The process of cell division in a bacterium. (B) These two cells of the bacterium *Pseudomonas aeruginosa* have almost completed cytokinesis.

- **Segregation.** In eukaryotes, the newly replicated chromosomes are closely associated with each other (thus they are known as **sister chromatids**), and a mechanism called **mitosis** segregates them into two new nuclei.

- **Cytokinesis.** Cytokinesis proceeds differently in plant cells (which have a cell wall) than in animal cells (which do not).

The cells resulting from mitosis are identical to the parent cell in the amount and kind of DNA that they contain. This contrasts with the second mechanism of nuclear division, meiosis.

Meiosis is the process of nuclear division that occurs in cells involved with sexual reproduction. While the two products of mitosis are genetically identical to the cell that produced them—they both have the same DNA—the products of meiosis are not. As we will see in Section 11.5, meiosis generates diversity by shuffling the genetic material, resulting in new gene combinations. Meiosis plays a key role in the sexual life cycle.

11.1 RECAP

Four events are required for cell division: a reproductive signal, replication of the genetic material (DNA), segregation of replicated DNA, and separation of the two daughter cells (cytokinesis). In prokaryotes, cell division can be rapid; in eukaryotes, the process is more intricate, and the chromosomes must be duplicated before cell division can occur.

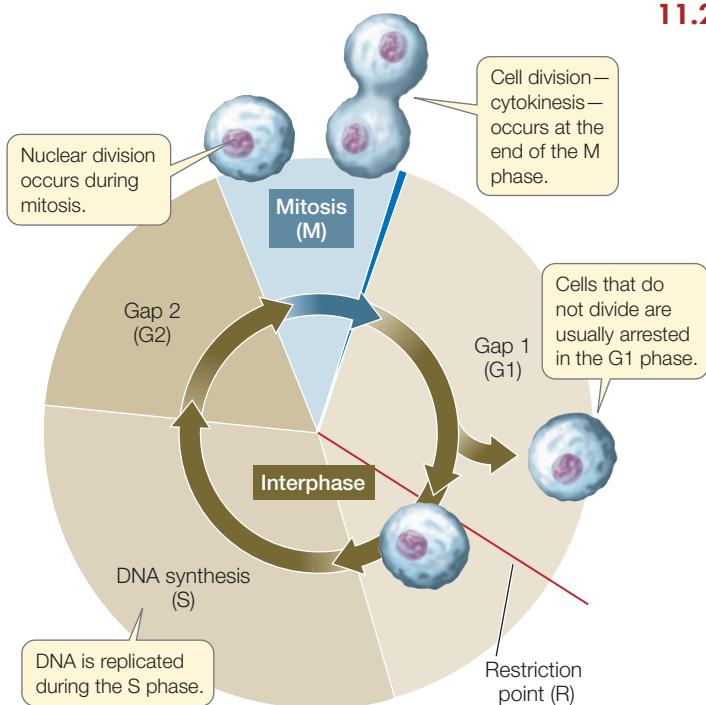
- What is the reproductive signal that leads the bacterium *Bacillus subtilis* to divide? **See p. 210**
- Explain why DNA must be replicated and segregated before a cell can divide. **See p. 210**
- What are the differences between cell division in prokaryotes (binary fission) and mitosis in eukaryotes? **See pp. 211–212**

What determines whether a cell will divide? How does mitosis lead to identical cells, and meiosis to diversity? Why do most eukaryotic organisms reproduce sexually? In the sections that follow, we will describe the details of mitosis and meiosis, and discuss their roles in development and evolution.

11.2 How Is Eukaryotic Cell Division Controlled?

As you will see throughout the book, different cells have different rates of cell division. Some cells, such as those in an early embryo, stem cells in bone marrow, or cells in the growing tip of a plant root, divide rapidly and continuously. Others, such as neurons in the brain or phloem cells in a plant stem, don't divide at all. Clearly, the signaling pathways for cells to divide are highly controlled.

The period between cell divisions is referred to as the **cell cycle**. The cell cycle can be divided into mitosis/cytokinesis and interphase. During **interphase**, the cell nucleus is visible and typical cell functions occur, including DNA replication. This phase of



11.3 The Eukaryotic Cell Cycle The cell cycle consists of a mitotic (M) phase, during which mitosis and cytokinesis take place, and a long period of growth known as interphase. Interphase has three subphases (G₁, S, and G₂) in cells that divide.

the cell cycle begins when cytokinesis is completed and ends when mitosis begins (Figure 11.3). In this section we will describe the events of interphase, especially those that trigger mitosis.

Cells, even when rapidly dividing, spend most of their time in interphase. So if we take a snapshot through the microscope of a cell population, most of the cells will be in interphase; only a small percentage will be in mitosis or cytokinesis at any given moment.

Interphase has three subphases, called G₁, S, and G₂. The cell's DNA replicates during **S phase** (the S stands for synthesis) (see Figure 11.3). The period between the end of cytokinesis and the onset of S phase is called **G₁**, or Gap 1. Another gap phase—**G₂**—separates the end of S phase and the beginning of mitosis. Mitosis and cytokinesis are referred to as the **M phase** of the cell cycle.

Let's look at the events of interphase in more detail:

- **G₁ phase.** During G₁, a cell is preparing for S phase, so at this stage each chromosome is a single, unreplicated structure. G₁ is quite variable in length in different cell types. Some rapidly dividing embryonic cells dispense with it entirely, while other cells may remain in G₁ for weeks or even years. In many cases these cells enter a resting phase called G₀. Special internal and external signals are needed to prompt a cell to leave G₀ and reenter the cell cycle at G₁.
- **The G₁-to-S transition.** At the G₁-to-S transition, called the **restriction point (R)**, the commitment is made to DNA replication and subsequent cell division (and thus another cell cycle).
- **S phase.** DNA replication occurs during S phase (see Section 13.3 for a detailed description). Each chromosome is duplicated and thereafter consists of two sister chromatids joined together and awaiting segregation into two new cells.

- **G₂ phase.** During G₂, the cell makes preparations for mitosis—for example, by synthesizing components of the microtubules that will move the chromatids to opposite ends of the dividing cell.

The initiation, termination, and operations of these phases are regulated by specific signals.

Specific signals trigger events in the cell cycle

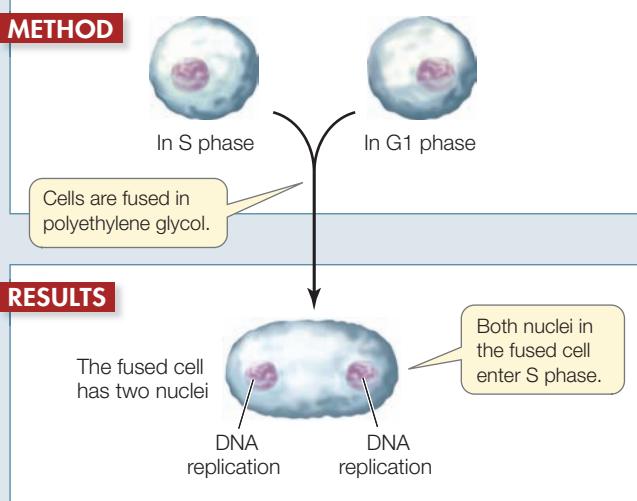
What events cause a cell to enter the S or M phases? A first indication that there were substances that control these transitions came from experiments involving *cell fusion*. Polyethylene glycol can be used to make different cells fuse together. Membrane lipids tend to partially dissolve in this nonpolar solvent, so that when it is present, cells will fuse their plasma membranes. Experiments involving the fusion of mammalian cells at different phases of the cell cycle showed that a cell in S phase produces a substance that activates DNA replication (Figure 11.4).

INVESTIGATING LIFE

11.4 Regulation of the Cell Cycle

Nuclei in G₁ do not undergo DNA replication, but nuclei in S phase do. To determine if there is some signal in the S cells that stimulates G₁ cells to replicate their DNA, cells in G₁ and S phases were fused together, creating cells with both G₁ and S properties.

HYPOTHESIS A cell in S phase contains an activator of DNA replication.



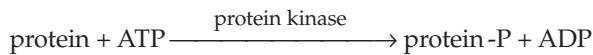
CONCLUSION The S phase cell produces a substance that diffuses to the G₁ nucleus and activates DNA replication.

FURTHER INVESTIGATION: How would you use this method to show that a cell in M phase produces an activator of mitosis?

Go to yourBioPortal.com for original citations, discussions, and relevant links for all INVESTIGATING LIFE figures.

Similar experiments point to a molecular activator for entry into M phase. As you will see, the signals that control progress through the cell cycle act through protein kinases.

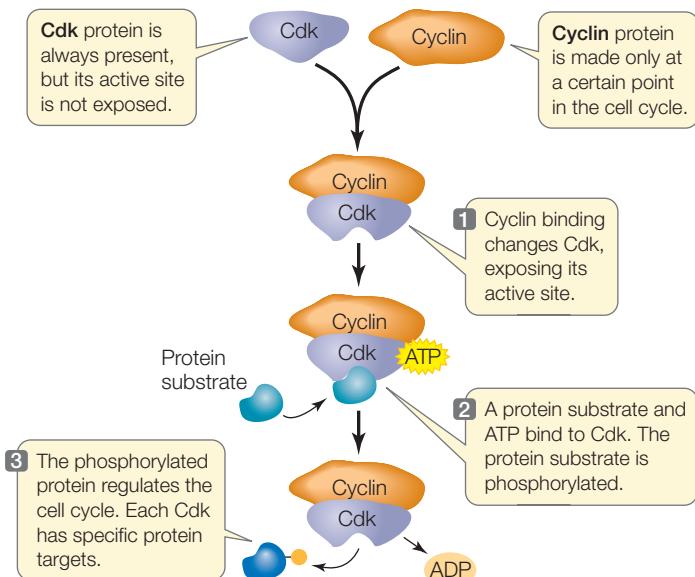
Progress through the cell cycle depends on the activities of **cyclin-dependent kinases**, or **Cdk's**. Recall from Section 7.1 that a *protein kinase* is an enzyme that catalyzes the transfer of a phosphate group from ATP to a target protein; this phosphate transfer is called *phosphorylation*.



By catalyzing the phosphorylation of certain target proteins, Cdk's play important roles at various points in the cell cycle. The discovery that Cdk's induce cell division is a beautiful example of how research on different organisms and different cell types can converge on a single mechanism. One group of scientists, led by James Maller at the University of Colorado, was studying immature sea urchin eggs, trying to find out how they are stimulated to divide and eventually form a mature egg. A protein called *matured promoting factor* was purified from maturing eggs, which by itself prodded immature egg cells to divide.

Meanwhile, Leland Hartwell at the University of Washington was studying the cell cycle in yeast (a single-celled eukaryote, see Figure 11.1A), and found a strain that was stalled at the G1-S boundary because it lacked a Cdk. It turned out that this yeast Cdk and the sea urchin maturation promoting factor had similar properties, and further work confirmed that the sea urchin protein was indeed a Cdk. Similar Cdk's were soon found to control the G1-to-S transition in many other organisms, including humans. Then others were found to control other parts of the cell cycle.

Cdk's are not active by themselves. As their name implies, cyclin-dependent kinases need to be activated by binding to a

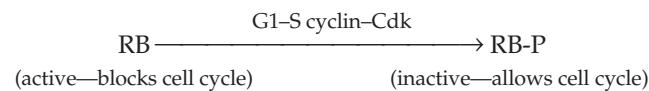


11.5 Cyclin Binding Activates Cdk Binding of a cyclin changes the three-dimensional structure of an inactive Cdk, making it an active protein kinase. Each cyclin-Cdk complex phosphorylates a specific target protein in the cell cycle.

second type of protein, called **cyclin**. This binding—an example of *allosteric regulation* (see Section 8.5)—activates the Cdk by altering its shape and exposing its active site (Figure 11.5).

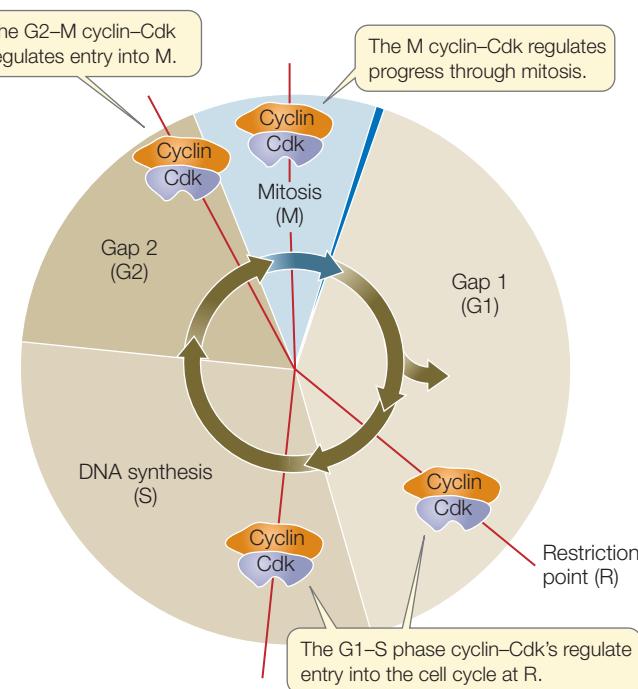
The cyclin-Cdk that controls passage from G1 to S phase is not the only such complex involved in regulating the eukaryotic cell cycle. There are different cyclin-Cdk's that act at different stages of the cycle (Figure 11.6). Let's take a closer look at G1-S cyclin-Cdk, which was the first to be discovered.

G1-S cyclin-Cdk catalyzes the phosphorylation of a protein called *retinoblastoma protein* (RB). In many cells, RB or a protein like it acts as an inhibitor of the cell cycle at the R (for “restriction”) point in late G1. To begin S phase, a cell must get by the RB block. Here is where G1-S cyclin-Cdk comes in: it catalyzes the addition of a phosphate to RB. This causes a change in the three-dimensional structure of RB, thereby inactivating it. With RB out of the way, the cell cycle can proceed. To summarize:

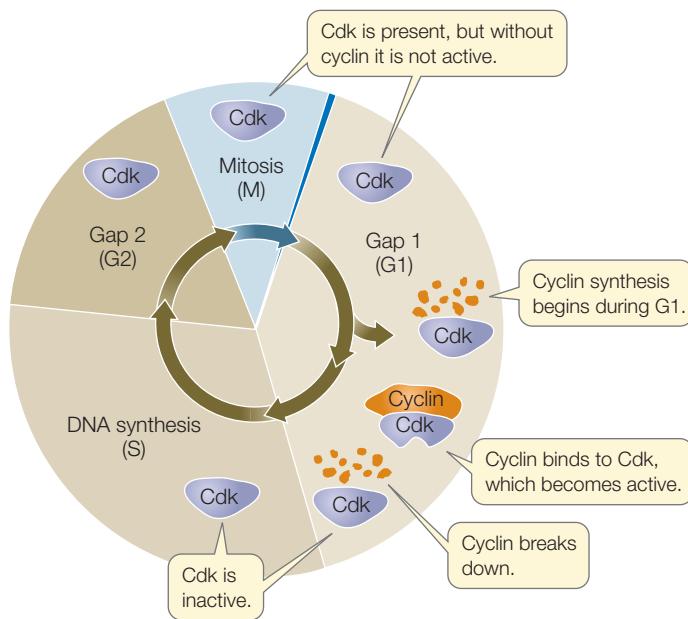


Progress through the cell cycle is regulated by the activities of Cdk's, and so regulating *them* is a key to regulating cell division. An effective way to regulate Cdk's is to regulate the presence or absence of cyclins (Figure 11.7). Simply put, if a cyclin is not present, its partner Cdk is not active. As their name suggests, the presence of cyclins is cyclic: they are made only at certain times of in the cell cycle.

The different cyclin-Cdk's act at **cell cycle checkpoints**, points at which a cell cycle's progress is regulated. For example, if a cell's DNA is substantially damaged by radiation or



11.6 Cyclin-Dependent Kinases Regulate Progress Through the Cell Cycle By acting at checkpoints (red lines), different cyclin-Cdk complexes regulate the orderly sequence of events in the cell cycle.



11.7 Cyclins Are Transient in the Cell Cycle Cyclins are made at a particular time and then break down. In this case, the cyclin is present during G₁ and activates a Cdk at that time.

toxic chemicals, it may be prevented from successfully completing a cell cycle. For DNA damage, there are three checkpoints:

- During G₁, before the cell enters S phase (restriction point)
- During S phase
- After S phase, during G₂

Let's consider the G₁ checkpoint. If DNA is damaged by radiation during G₁, a protein called p21 is made. (The *p* stands for "protein" and the 21 stands for its molecular weight—about 21,000 daltons.) The p21 protein can bind to the G₁-S Cdk, preventing its activation by cyclin. So the cell cycle stops while repairs are made to the DNA (you will learn more about DNA repair in Section 13.4). The p21 protein breaks down after the DNA is repaired, allowing cyclin to bind to the Cdk so that the cell cycle can proceed. If DNA damage is severe and it cannot be repaired, the cell will undergo programmed cell death (apoptosis, which we will discuss in Section 11.6).

In addition to these internal signals, the cell cycle is influenced by signals from the extracellular environment.

Growth factors can stimulate cells to divide

Cyclin-Cdk's provide cells with internal controls of their progress through the cell cycle. Not all cells in an organism go through the cell cycle on a regular basis. Some cells either no longer go through the cell cycle and enter G₀, or go through it slowly and divide infrequently. If such cells are to divide, they must be stimulated by external chemical signals called **growth factors**. These proteins activate a signal transduction pathway that often ends up with the activation of Cdk's (signal transduction is discussed in Chapter 7):

- If you cut yourself and bleed, specialized cell fragments called *platelets* gather at the wound to initiate blood clotting. The platelets produce and release a protein called *platelet-*

derived growth factor that diffuses to the adjacent cells in the skin and stimulates them to divide and heal the wound.

- Red and white blood cells have limited lifetimes and must be replaced through the division of immature, unspecialized blood cell precursors in the bone marrow. Two types of growth factors, *interleukins* and *erythropoietin*, stimulate the division and specialization, respectively, of precursor cells.

In these and other examples, growth factors bind to specific receptors on target cells, and activate signal transduction pathways the end with cyclin synthesis, thereby activating Cdk's and the cell cycle. As you can see from the examples, growth factors are important in maintaining homeostasis.

11.2 RECAP

The eukaryotic cell cycle is under both external and internal control. Cdk's control the eukaryotic cell cycle and are themselves controlled by cyclins. External signals such as growth factors can initiate the cell cycle.

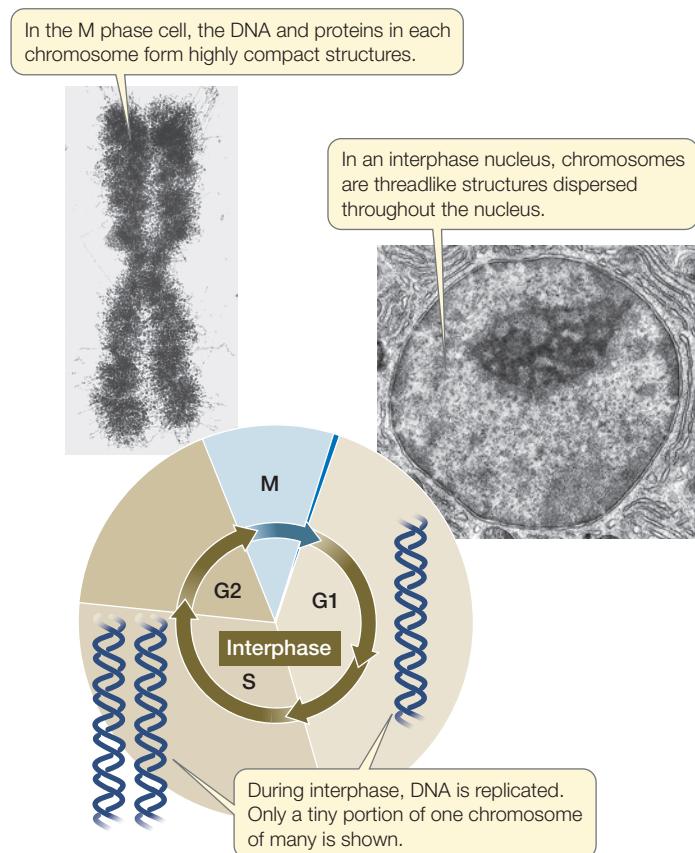
- Draw a cell cycle diagram showing the various stages of interphase. See pp. 212–213 and Figure 11.3
- How do cyclin-Cdk's control the progress of the cell cycle? See pp. 214–215 and Figure 11.6
- What are the differences between external and internal controls of the cell cycle? See p. 215

11.3 What Happens during Mitosis?

The third essential step in the process of cell division—segregation of the replicated DNA—occurs during mitosis. Prior to segregation, the huge DNA molecules and their associated proteins in each chromosome become condensed into more compact structures. After segregation by mitosis, cytokinesis separates the two cells. Let's now look at these steps more closely.

Prior to mitosis, eukaryotic DNA is packed into very compact chromosomes

A eukaryotic chromosome consists of one or two gigantic, linear, double-stranded DNA molecules complexed with many proteins (the complex of DNA and proteins is referred to as **chromatin**). Before S phase, each chromosome contains only one double-stranded DNA molecule. After it replicates during S phase, however, there are two double-stranded DNA molecules, known as sister chromatids. The sister chromatids are held together along most of their length by a protein complex called **cohesin**. They stay this way throughout interphase G₂ until mitosis, when most of the cohesin is removed, except in a region called the **centromere** at which the chromatids remain held together. At the end of G₂, a second group of proteins called **condensins** coat the DNA molecules and makes them more compact (Figure 11.8).



11.8 Chromosomes, Chromatids, and Chromatin

DNA in the interphase nucleus is diffuse and becomes compacted as mitosis begins.

If all the DNA in a typical human cell were put end to end, it would be nearly 2 meters long. Yet the nucleus is only 5 μm (0.000005 meters) in diameter. So eukaryotic DNA, like that in prokaryotes, is extensively packaged in a highly organized way (**Figure 11.9**). This packing is achieved largely by proteins that are closely associated with the DNA; these proteins are called **histones** (*histos*, “web” or “loom”). They are positively charged at cellular pH levels because of their high content of the basic amino acids lysine and arginine. These positive charges attract the negative phosphate groups on DNA. These DNA–histone interactions, as well as histone–histone interactions, result in the formation of beadlike units called **nucleosomes**.

During interphase, the chromatin that makes up each chromosome is much less densely packaged, and consists of single DNA molecules running around vast numbers of nucleosomes like beads on a string. During this phase of the cell cycle, the DNA is accessible to proteins involved in replication and transcription. Once a mitotic chromosome is formed its compact nature makes it inaccessible to replication and transcription factors, and so these processes cannot occur.

During the early stages of both mitosis and meiosis, the chromatin becomes ever more tightly coiled and condensed as the nucleosomes pack together. Further coiling of the chromatin continues up to the time at which the chromatids begin to move apart.

Overview: Mitosis segregates copies of genetic information

In mitosis, a single nucleus gives rise to two nuclei that are genetically identical to each other and to the parent nucleus. Mitosis (the M phase of the cell cycle) ensures the accurate segregation of the eukaryotic cell’s multiple chromosomes into the daughter nuclei. While mitosis is a continuous process in which each event flows smoothly into the next, it is convenient to subdivide it into a series of stages: prophase, prometaphase, metaphase, anaphase, and telophase. Before we consider each of these stages, we will describe two cellular structures that contribute to the orderly segregation of the chromosomes during mitosis—the centrosome and the spindle.

The centrosomes determine the plane of cell division

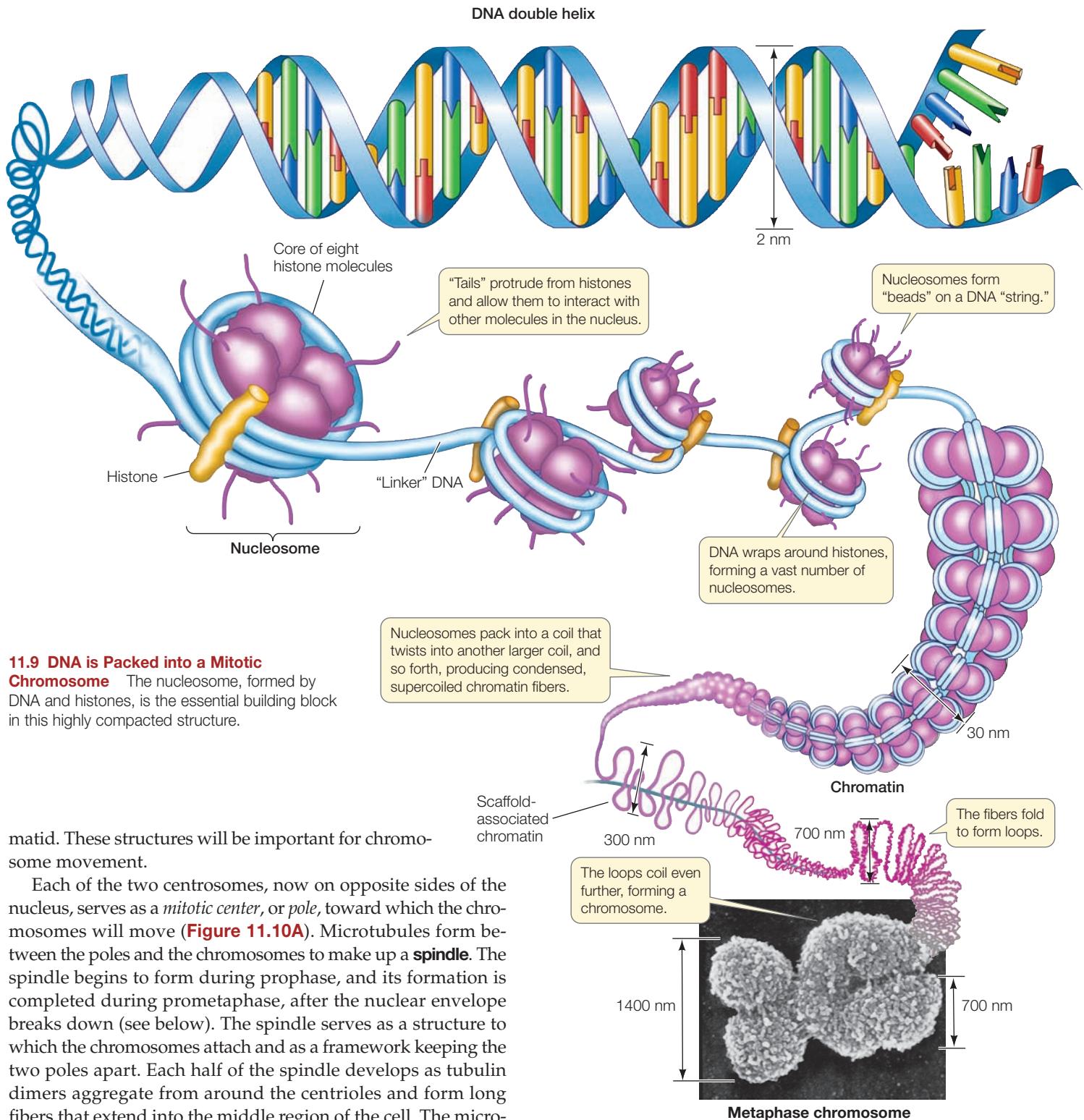
Before the spindle apparatus for chromosome segregation forms, the orientation of this spindle is determined. This is accomplished by the **centrosome** (“central body”), an organelle in the cytoplasm near the nucleus. In many organisms, each centrosome consists of a pair of **centrioles**, each one a hollow tube formed by nine triplets of microtubules. The two tubes are at right angles to each other.

During S phase the centrosome doubles to form a pair of centrosomes. At the G2-to-M transition, the two centrosomes separate from one another, moving to opposite ends of the nuclear envelope. Eventually these will identify “poles” toward which chromosomes will move during segregation. The positions of the centrosomes determine the plane at which the cell will divide; therefore they determine the spatial relationship between the two new cells. This relationship may be of little consequence to single free-living cells such as yeasts, but it is important for cells in a multicellular organism. For example, during development from a fertilized egg to an embryo, the daughter cells from some divisions must be positioned correctly to receive signals to form new tissues.

The centrioles are surrounded by high concentrations of tubulin dimers, and these proteins aggregate to form the microtubules that orchestrate chromosomal movement. (Plant cells lack centrosomes, but distinct microtubule organizing centers at each end of the cell play the same role.) These microtubules are the major part of the spindle structure, which is required for the orderly segregation of the chromosomes.

The spindle begins to form during prophase

During interphase, only the nuclear envelope, the nucleoli (see Section 5.3), and a barely discernible tangle of chromatin are visible under the light microscope. The appearance of the nucleus changes as the cell enters **prophase**—the beginning of mitosis. Most of the cohesin that has held the two products of DNA replication together since S phase is removed, so the individual chromatids become visible. They are still held together by a small amount of cohesin at the centromere. Late in prophase, specialized three-layered structures called **kinetochores** develop in the centromere region, one on each chro-



11.9 DNA is Packed into a Mitotic Chromosome

Chromatin The nucleosome, formed by DNA and histones, is the essential building block in this highly compacted structure.

matid. These structures will be important for chromosome movement.

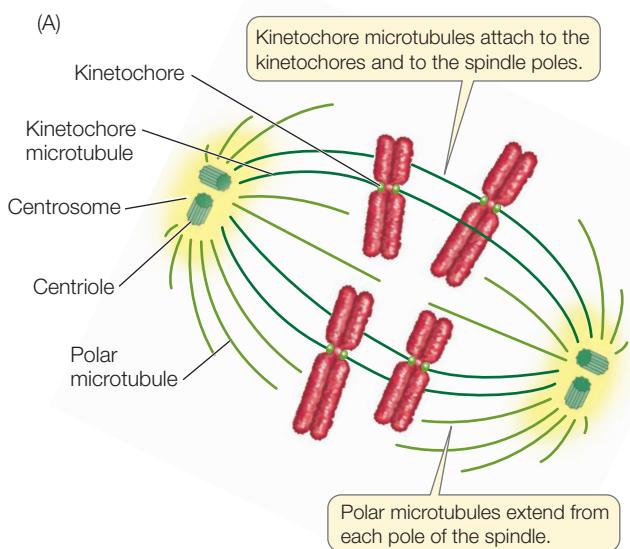
Each of the two centrosomes, now on opposite sides of the nucleus, serves as a *mitotic center*, or *pole*, toward which the chromosomes will move (Figure 11.10A). Microtubules form between the poles and the chromosomes to make up a **spindle**. The spindle begins to form during prophase, and its formation is completed during prometaphase, after the nuclear envelope breaks down (see below). The spindle serves as a structure to which the chromosomes attach and as a framework keeping the two poles apart. Each half of the spindle develops as tubulin dimers aggregate from around the centrioles and form long fibers that extend into the middle region of the cell. The microtubules are initially unstable, constantly forming and falling apart, until they contact kinetochores or microtubules from the other half-spindle and become more stable.

There are two types of microtubule in the spindle:

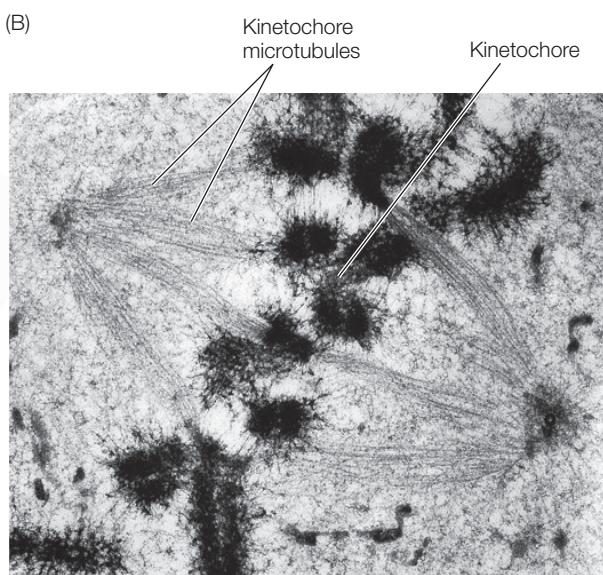
- **Polar microtubules** form the framework of the spindle, and run from one pole to the other.
- **Kinetochores microtubules**, which form later, attach to the kinetochores on the chromosomes. The two sister chromatids in each chromosome pair become attached to kine-

tochore microtubules in opposite halves of the spindle (Figure 11.10B). This ensures that the two chromatids will eventually move to opposite poles.

Movement of the chromatids is the central feature of mitosis. It accomplishes the segregation that is needed for cell division and completion of the cell cycle. Prophase prepares for this movement, and the actual segregation takes place in the next three phases of mitosis.

**11.10 The Mitotic Spindle Consists of Microtubules**

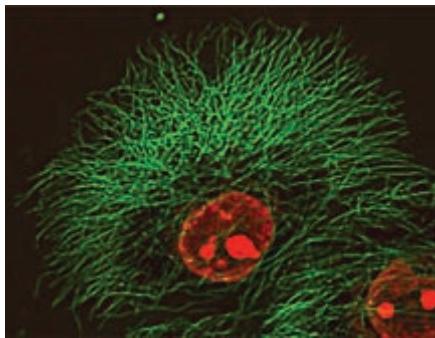
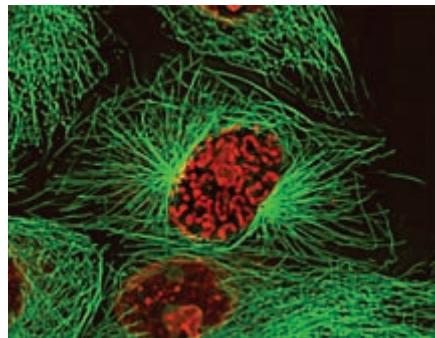
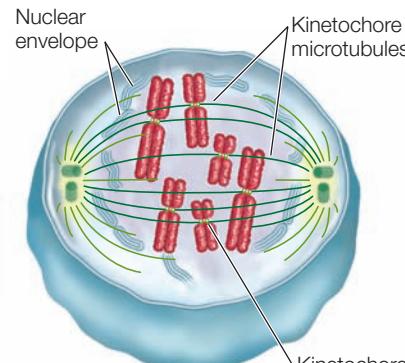
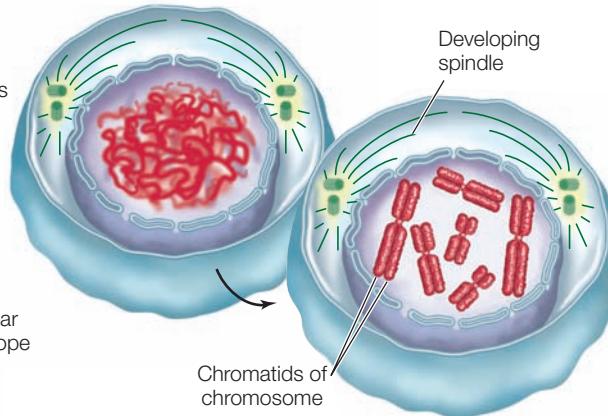
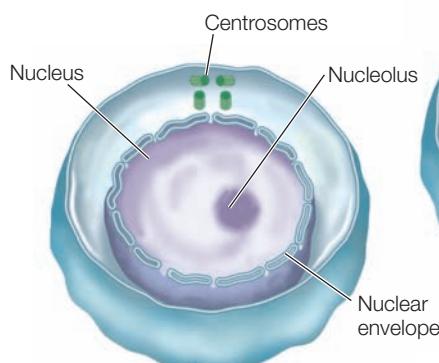
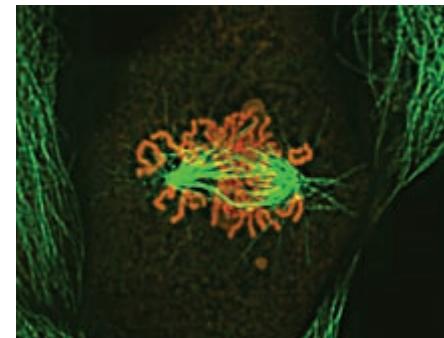
(A) The spindle apparatus in an animal cell at metaphase. In plant cells, centrioles are not present. (B) An electron micrograph of metaphase emphasizing the kinetochore microtubules.



yourBioPortal.com

[GO TO Web Activity 11.1 • The Mitotic Spindle](#)

11.11 The Phases of Mitosis Mitosis results in two new nuclei that are genetically identical to each other and to the nucleus from which they were formed. In the micrographs, the green dye stains microtubules (and thus the spindle); the red dye stains the chromosomes. The chromosomes in the diagrams are stylized to emphasize the fates of the individual chromatids.

Interphase**Prophase****Prometaphase**

1 During the S phase of interphase, the nucleus replicates its DNA and centrosomes.

2 The chromatin coils and supercoils, becoming more and more compact and condensing into visible chromosomes. The chromosomes consist of identical, paired sister chromatids. Centrosomes move to opposite poles.

3 The nuclear envelope breaks down. Kinetochore microtubules appear and connect the kinetochores to the poles.

yourBioPortal.com

[GO TO Web Activity 11.2 • Images of Mitosis](#)

Chromosome separation and movement are highly organized

During the next three phases, prometaphase, metaphase, and anaphase, dramatic changes take place in the cell and the chromosomes.

- **Prometaphase.** The nuclear envelope breaks down and the compacted chromosomes consisting of two chromatids attach to the kinetochore microtubules.
- **Metaphase.** The chromosomes line up at the midline of the cell (equatorial position).
- **Anaphase.** The chromatids separate and move away from each other toward the poles.

You will find these events depicted and described in **Figure 11.11**. Here, we will consider two key processes: separation of the chromatids, and the mechanism of their actual movement toward the poles.

CHROMATID SEPARATION The separation of chromatids occurs at the beginning of anaphase. It is controlled by M phase cyclin-Cdk (see Figure 11.6), which activates another protein complex called the *anaphase-promoting complex* (APC). Separation occurs because one subunit of the cohesin protein holding the sister chromatids together is hydrolyzed by a specific protease,

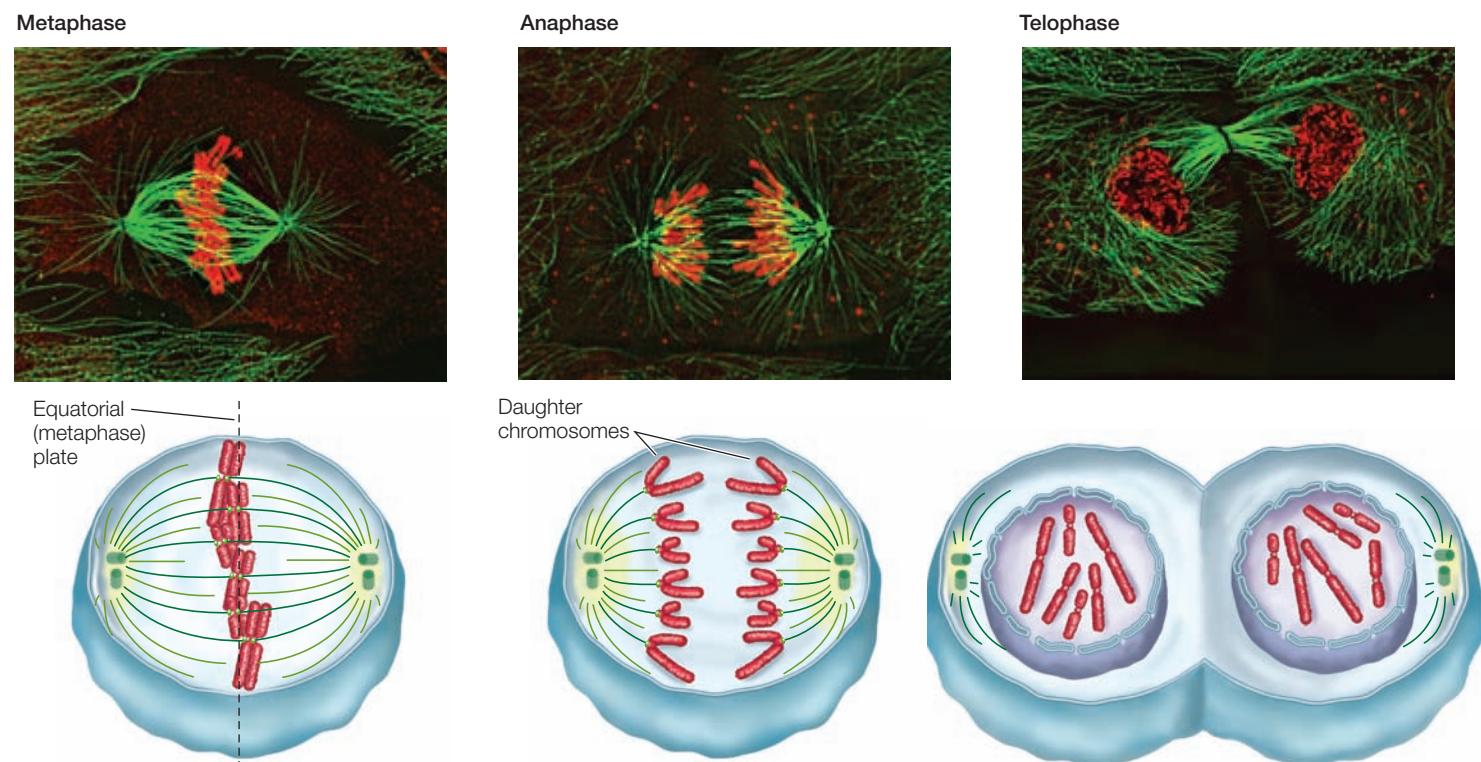
appropriately called *separase* (**Figure 11.12**). After they separate, the chromatids are called **daughter chromosomes**.

CHROMOSOME MOVEMENT The migration of the two sets of daughter chromosomes to the poles of the cell is a highly organized, active process. Two mechanisms operate to move the chromosomes along. First, the kinetochores contain a protein called *cytoplasmic dynein* that acts as a “molecular motor.” It hydrolyzes ATP to ADP and phosphate, thus releasing energy that may move the chromosomes along the microtubules toward the poles. This accounts for about 75 percent of the force of motion. Second, the kinetochore microtubules shorten from the poles, drawing the chromosomes toward them, accounting for about 25 percent of the force of motion.

- **Telophase** occurs after the chromosomes have separated and is the last phase of mitosis. During this period, a nuclear envelope forms around each set of chromosomes, nucleoli appear, and the chromosomes become less compact. The spindle also disappears at this time. As a result, there are two new nuclei in a single cell.

Cytokinesis is the division of the cytoplasm

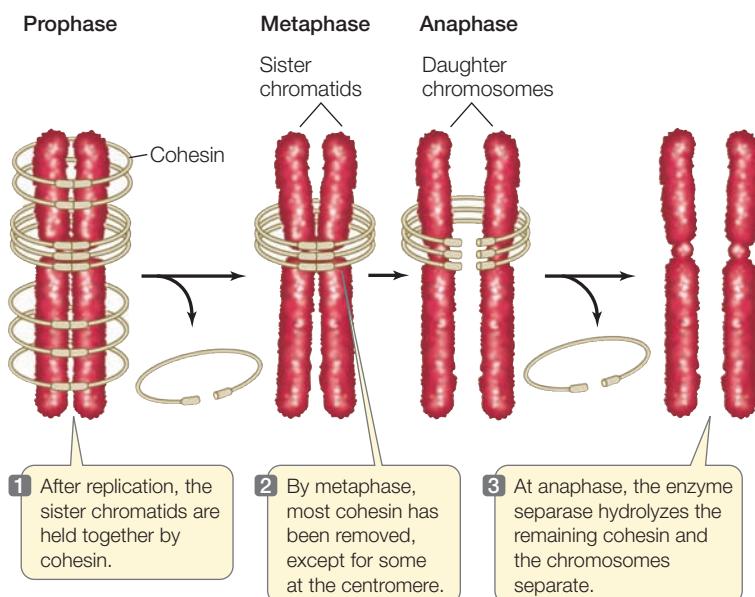
Mitosis refers only to the division of the nucleus. The division of the cell’s cytoplasm, which follows mitosis, is called cytoki-



4 The centromeres become aligned in a plane at the cell's equator.

5 The paired sister chromatids separate, and the new daughter chromosomes begin to move toward the poles.

6 The daughter chromosomes reach the poles. As telophase concludes, the nuclear envelopes and nucleoli re-form, the chromatin decondenses, and, after cytokinesis, the daughter cells enter interphase once again.



11.12 Chromatid Attachment and Separation The cohesin protein complex holds sister chromatids together at the centromere. The enzyme separase hydrolyzes cohesin at the onset of anaphase, allowing the chromatids to separate into daughter chromosomes.

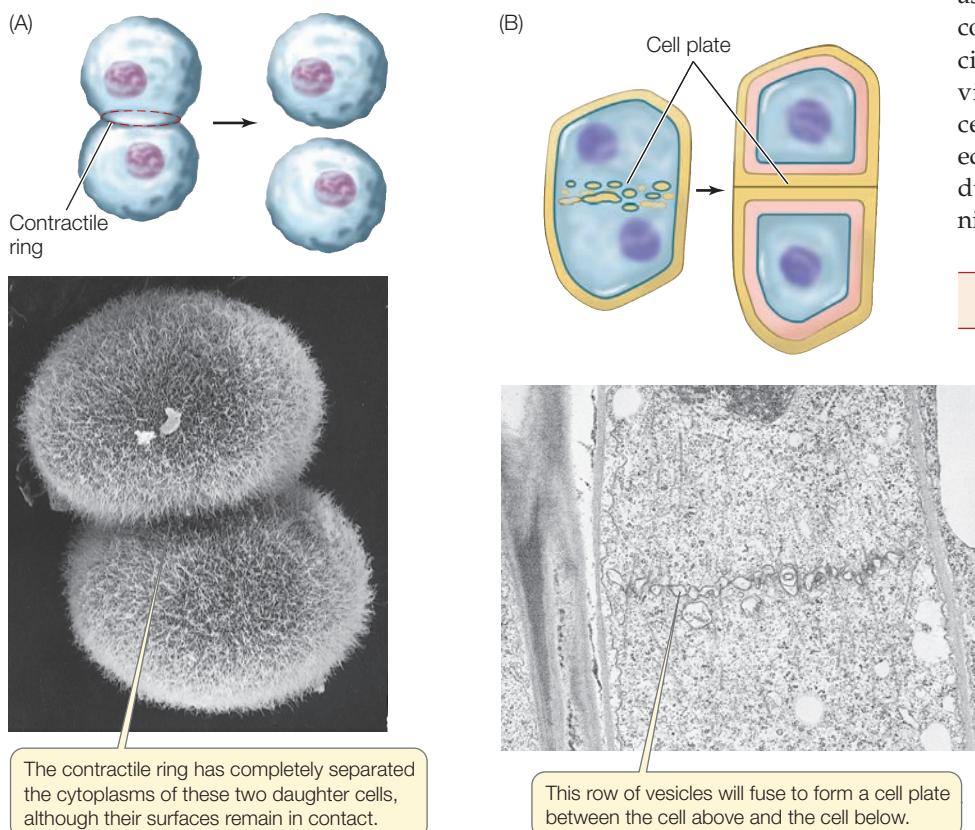
mic surface of the plasma membrane. These two proteins interact to produce a contraction, just as they do in muscles, thus pinching the cell in two. The microfilaments assemble rapidly from actin monomers that are present in the interphase cytoskeleton. Their assembly is under the control of calcium ions that are released from storage sites in the center of the cell.

The plant cell cytoplasm divides differently because plants have cell walls that are rigid. In plant cells, as the spindle breaks down after mitosis, membranous vesicles derived from the Golgi apparatus appear along the plane of cell division, roughly midway between the two daughter nuclei. The vesicles are propelled along microtubules by the motor protein kinesin, and fuse to form a new plasma membrane. At the same time they contribute their contents to a *cell plate*, which is the beginning of a new cell wall (**Figure 11.13B**).

Following cytokinesis, each daughter cell contains all the components of a complete cell. A precise distribution of chromosomes is ensured by mitosis. In contrast, organelles such as ribosomes, mitochondria, and chloroplasts need not be distributed equally between daughter cells as long as some of each are present in each cell. Accordingly, there is no mechanism with a precision comparable to that of mitosis to provide for their equal allocation to daughter cells. As we will see in Chapter 19, the unequal distribution of cytoplasmic components during development can have functional significance for the two new cells.

yourBioPortal.com

GO TO Animated Tutorial 11.1 • Mitosis



11.13 Cytokinesis Differs in Animal and Plant Cells (A) A sea urchin zygote (fertilized egg) that has just completed cytokinesis at the end of the first cell division of its development into an embryo. (B) A dividing plant cell in late telophase. Plant cells divide differently from animal cells because plant cells have cell walls.

11.3 RECAP

Mitosis is the division of the nucleus of a eukaryotic cell into two nuclei identical to each other and to the parent nucleus. The process of mitosis, while continuous, can be viewed as a series of events (prophase, prometaphase, metaphase, anaphase, and telophase). Once two identical nuclei have formed, the cell divides into two cells by cytokinesis.

- What is the difference between a chromosome, a chromatid, and a daughter chromosome? **See Figures 11.8 and 11.11**
- What are the various levels of “packing” by which the genetic information contained in linear DNA is condensed during prophase? **See p. 216 and Figure 11.9**
- Describe how chromosomes move during mitosis. **See p. 219 and Figure 11.11**
- What are the differences in cytokinesis between plant and animal cells? **See p. 220 and Figure 11.13**

The intricate process of mitosis results in two cells that are genetically identical. But, as mentioned earlier, there is another eukaryotic cell division process, called meiosis, that results in genetic diversity. What is the role of that process?

11.4 What Role Does Cell Division Play in a Sexual Life Cycle?

The mitotic cell cycle repeats itself and by this process, a single cell can give rise to a vast number of cells with identical nuclear DNA. Meiosis, on the other hand, produces just four daughter

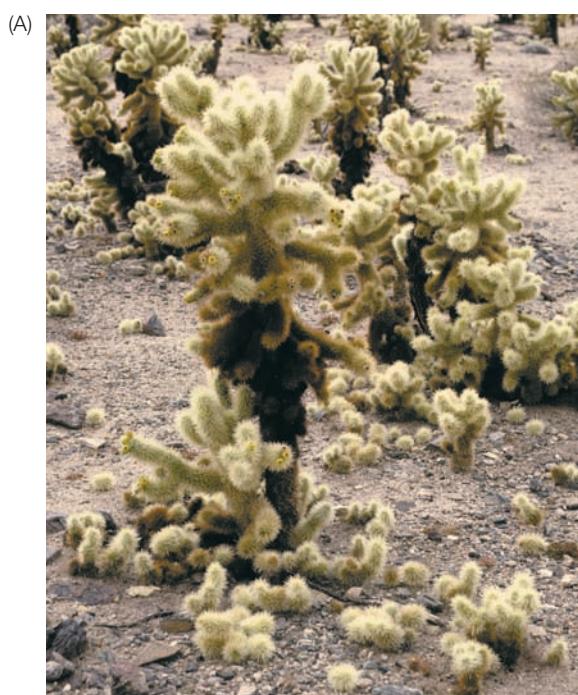
cells. Mitosis and meiosis are both involved in reproduction, but they have different roles: asexual reproduction involves only mitosis, while sexual reproduction involves both mitosis and meiosis.

Asexual reproduction by mitosis results in genetic constancy

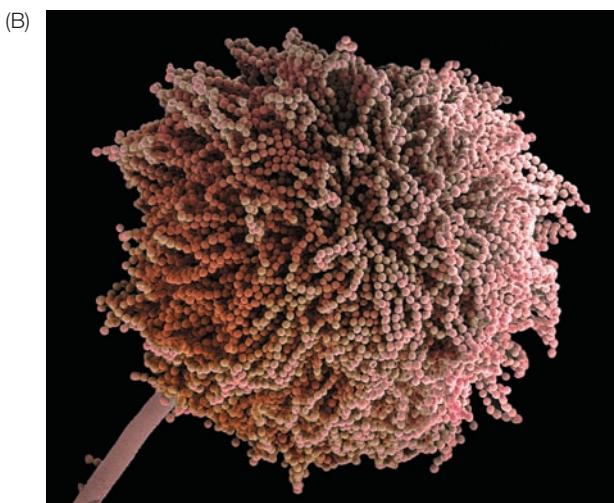
Asexual reproduction, sometimes called *vegetative reproduction*, is based on the mitotic division of the nucleus. An organism that reproduces asexually may be single-celled like yeast, reproducing itself with each cell cycle, or it may be multicellular like the cholla cactus, that breaks off a piece to produce a new multicellular organism (**Figure 11.14**). Asexual reproduction is a rapid and effective means of making new individuals, and it is common in nature. In asexual reproduction, the offspring are **clones** of the parent organism; that is, the offspring are *genetically identical* to the parent. Any genetic variation among the offspring is most likely due to small environmentally caused changes in the DNA, called *mutations*. As you will see, this small amount of variation contrasts with the extensive variation possible in sexually reproducing organisms.

Sexual reproduction by meiosis results in genetic diversity

Unlike asexual reproduction, **sexual reproduction** results in an organism that is not identical to its parents. Sexual reproduction requires **gametes** created by meiosis; two parents each contribute one gamete to each of their offspring. Meiosis can produce gametes—and thus offspring—that differ genetically from each other and from the parents. Because of this genetic variation, some offspring may be better adapted than others to sur-



11.14 Asexual Reproduction in the Large and the Small (A) Some cacti like this cholla have brittle stems that break off easily. Fragments on the ground set down roots and develop by mitotic cell divisions into new plants that are genetically identical to the plant they came from. (B) These strings of cells are asexual spores formed by a fungus. Each spore contains a nucleus produced by a mitotic division and is genetically identical to the parent that produced it. It can divide to form a new fungus.



vive and reproduce in a particular environment. Meiosis thus generates the genetic diversity that is the raw material for natural selection and evolution.

In most multicellular organisms, the body cells that are *not* specialized for reproduction, called **somatic cells**, each contain two sets of chromosomes, which are found in pairs. One chromosome of each pair comes from each of the organism's two parents; for example, in humans with 46 chromosomes, 23 come from the mother and 23 from the father. The members of such a **homologous pair** are similar in size and appearance, except for the sex chromosomes found in some species (see Section 12.4). The two chromosomes in a homologous pair (called **homologs**) bear corresponding, though often not identical, genetic information. For example, a homologous pair of chromosomes in a plant may carry different versions of a gene that controls seed shape. One homolog may carry the version for wrinkled seeds while the other may carry the version for smooth seeds.

Gametes, on the other hand, contain only a single set of chromosomes—that is, one homolog from each pair. The number of chromosomes in a gamete is denoted by n , and the cell is said to be **haploid**. Two haploid gametes fuse to form a **zygote**, in a process called **fertilization**. The zygote thus has two sets of chromosomes, just as somatic cells do. Its chromosome number is denoted by $2n$, and the zygote is said to be **diploid**. Depending on the organism, the zygote may divide by either meiosis or mi-

tosis. Either way, a new mature organism develops that is capable of sexual reproduction.

All sexual life cycles involve meiosis to produce gametes or cells that are haploid. Eventually, the haploid cells or gametes fuse to produce a zygote, beginning the diploid stage of the life cycle. Since the origin of sexual reproduction, evolution has generated many different versions of the sexual life cycle. **Figure 11.15** presents three examples.

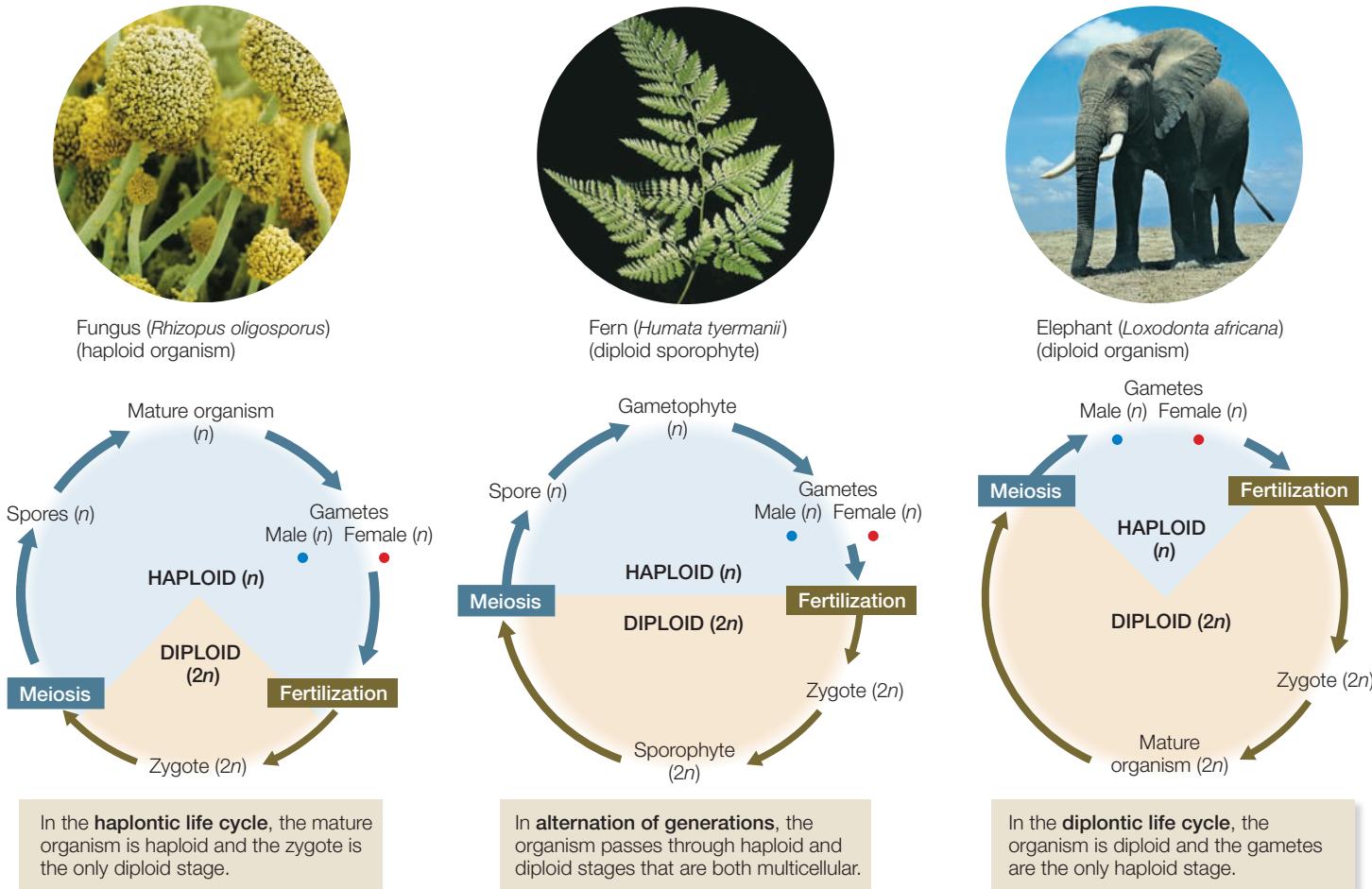
- In **haplontic** organisms, including most protists, fungi, and some green algae, the tiny zygote is the only diploid cell in the life cycle. After it is formed it immediately undergoes meiosis to produce more haploid cells. These are usually **spores**, which are the dispersal units for the organism, like the seeds of a plant. A spore germinates to form a new haploid organism, which may be single-celled or multicellular. Cells of the mature haploid organism fuse to form the diploid zygote.
- Most plants and some fungi display **alternation of generations**. As for many haplontic organisms, meiosis gives rise

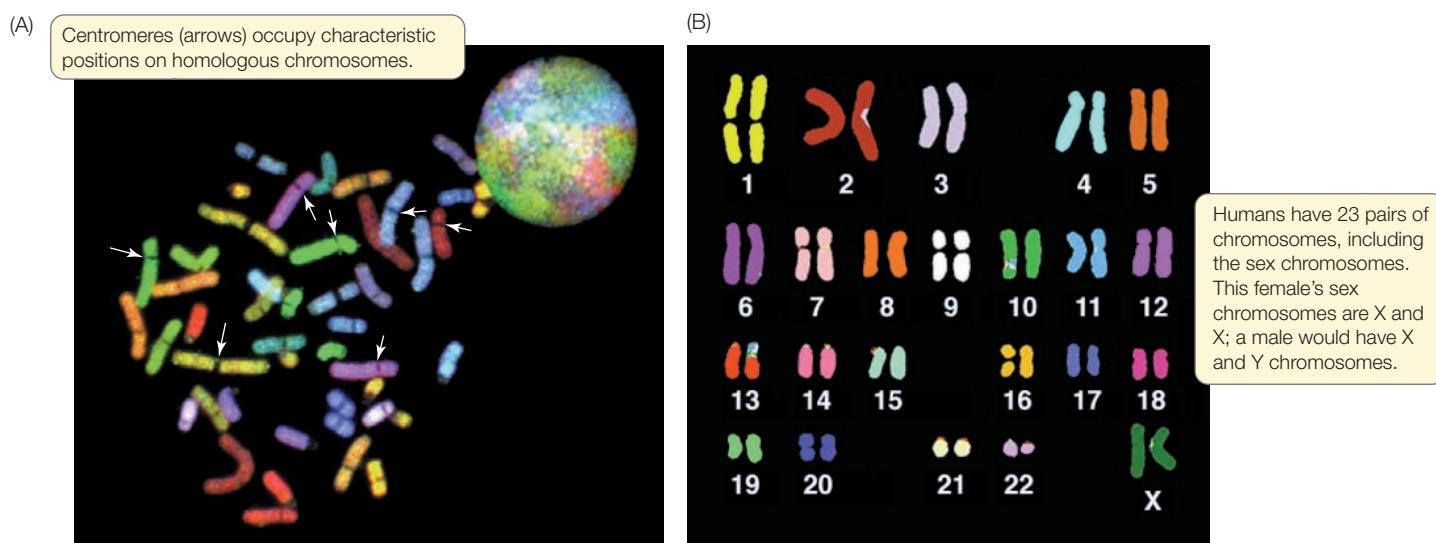
11.15 Fertilization and Meiosis Alternate in Sexual Reproduction

In sexual reproduction, haploid (n) cells or organisms alternate with diploid ($2n$) cells or organisms.

yourBioPortal.com

GO TO Web Activity 11.3 • Sexual Life Cycle





to haploid spores, which divide by mitosis to form a haploid life stage called the *gametophyte*. The gametophyte forms gametes by mitosis, which fuse to form a diploid zygote. The zygote divides by mitosis to become the diploid *sporophyte*, which in turn produces the gametes by meiosis.

- In *diplobiontic* organisms, which include animals, brown algae and some fungi, the gametes are the only haploid cells in the life cycle, and the mature organism is diploid.

These life cycles are described in greater detail in Part Seven. For now we will focus on the role of sexual reproduction in generating diversity among individual organisms.

The essence of sexual reproduction is the *random selection of half of the diploid chromosome set* to make a haploid gamete, followed by fusion of two haploid gametes to produce a diploid cell. Both of these steps contribute to a shuffling of genetic information in the population, so that no two individuals have exactly the same genetic constitution. The diversity provided by sexual reproduction opens up enormous opportunities for evolution.

The number, shapes, and sizes of the metaphase chromosomes constitute the karyotype

When cells are in metaphase of mitosis, it is often possible to count and characterize their individual chromosomes. If a photomicrograph of the entire set of chromosomes is made, the images of the individual chromosomes can be manipulated, pairing and placing them in an orderly arrangement. Such a rearranged photomicrograph reveals the number, shapes, and sizes of the chromosomes in a cell, which together constitute its **karyotype** (Figure 11.16). In humans, karyotypes can aid in the diagnosis of certain diseases, and this has led to an entire branch of medicine called *cytogenetics*. However, as you will see in Chapter 15, chromosome analysis with the microscope is being replaced by direct analysis of DNA.

Individual chromosomes can be recognized by their lengths, the positions of their centromeres, and characteristic banding

11.16 The Human Karyotype (A) Chromosomes from a human cell in metaphase. The DNA of each chromosome pair has a specific nucleotide sequence that is stained by a particular colored dye, so that the chromosomes in a homologous pair share a distinctive color. Each chromosome at this stage is composed of two chromatids, but they cannot be distinguished. At the upper right is an interphase nucleus. (B) This karyogram, produced by computerized analysis of the image on the left, shows homologous pairs lined up together and numbered, clearly revealing the individual's karyotype.

patterns that are visible when the chromosomes are stained and observed at high magnification. In diploid cells, the karyotype consists of homologous pairs of chromosomes—for example, there are 23 pairs and a total of 46 chromosomes in humans. There is no simple relationship between the size of an organism and its chromosome number. A housefly has 5 chromosome pairs and a horse has 32, but the smaller carp (a fish) has 52 pairs. Probably the highest number of chromosomes in any organism is in the fern *Ophioglossum reticulatum*, which has 1,260 (630 pairs)!

11.4 RECAP

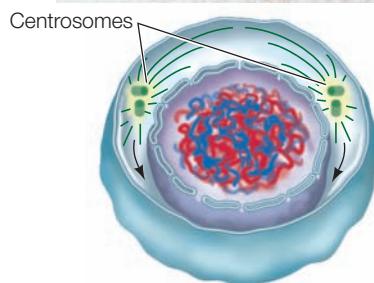
Meiosis is necessary for sexual reproduction, in which haploid gametes fuse to produce a diploid zygote. Sexual reproduction results in genetic diversity, the foundation of evolution.

- What is the difference, in terms of genetics, between asexual and sexual reproduction? **See p. 221**
- How does fertilization produce a diploid organism? **See p. 222**
- What general features do all sexual life cycles have in common? **See p. 222 and Figure 11.15**

Meiosis, unlike mitosis, results in daughter cells that have half as many chromosomes as the parent cell. Next we will look at the processes of meiosis.

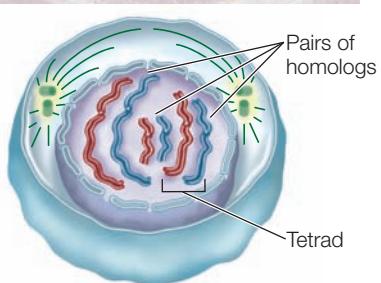
MEIOSIS I

Early prophase I



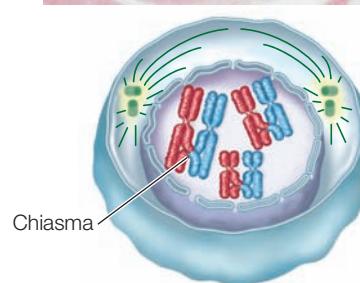
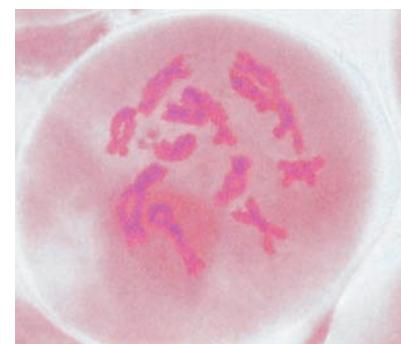
1 The chromatin begins to condense following interphase.

Mid-prophase I



2 Synapsis aligns homologs, and chromosomes condense further.

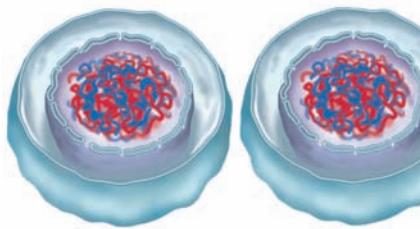
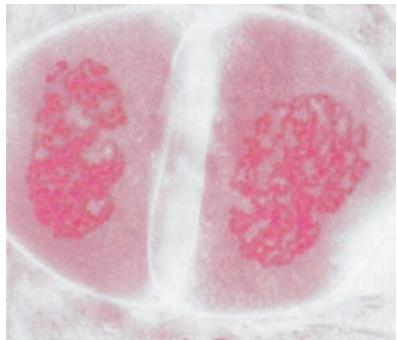
Late prophase I–Prometaphase



3 The chromosomes continue to coil and shorten. The chiasmata reflect crossing over, the exchange of genetic material between nonsister chromatids in a homologous pair. In prometaphase the nuclear envelope breaks down.

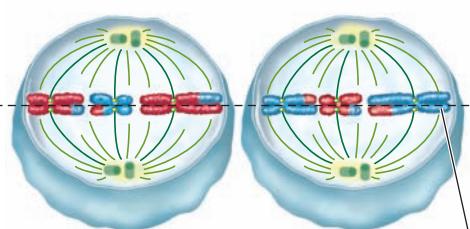
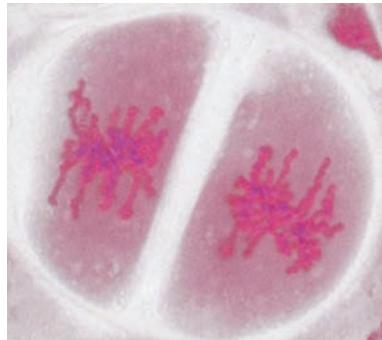
MEIOSIS II

Prophase II



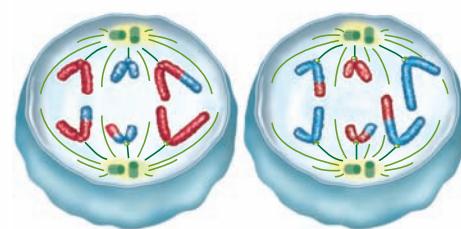
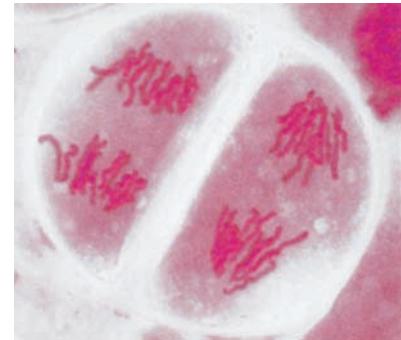
7 The chromosomes condense again, following a brief interphase (interkinesis) in which DNA does not replicate.

Metaphase II



8 The centromeres of the paired chromatids line up across the equatorial plates of each cell.

Anaphase II



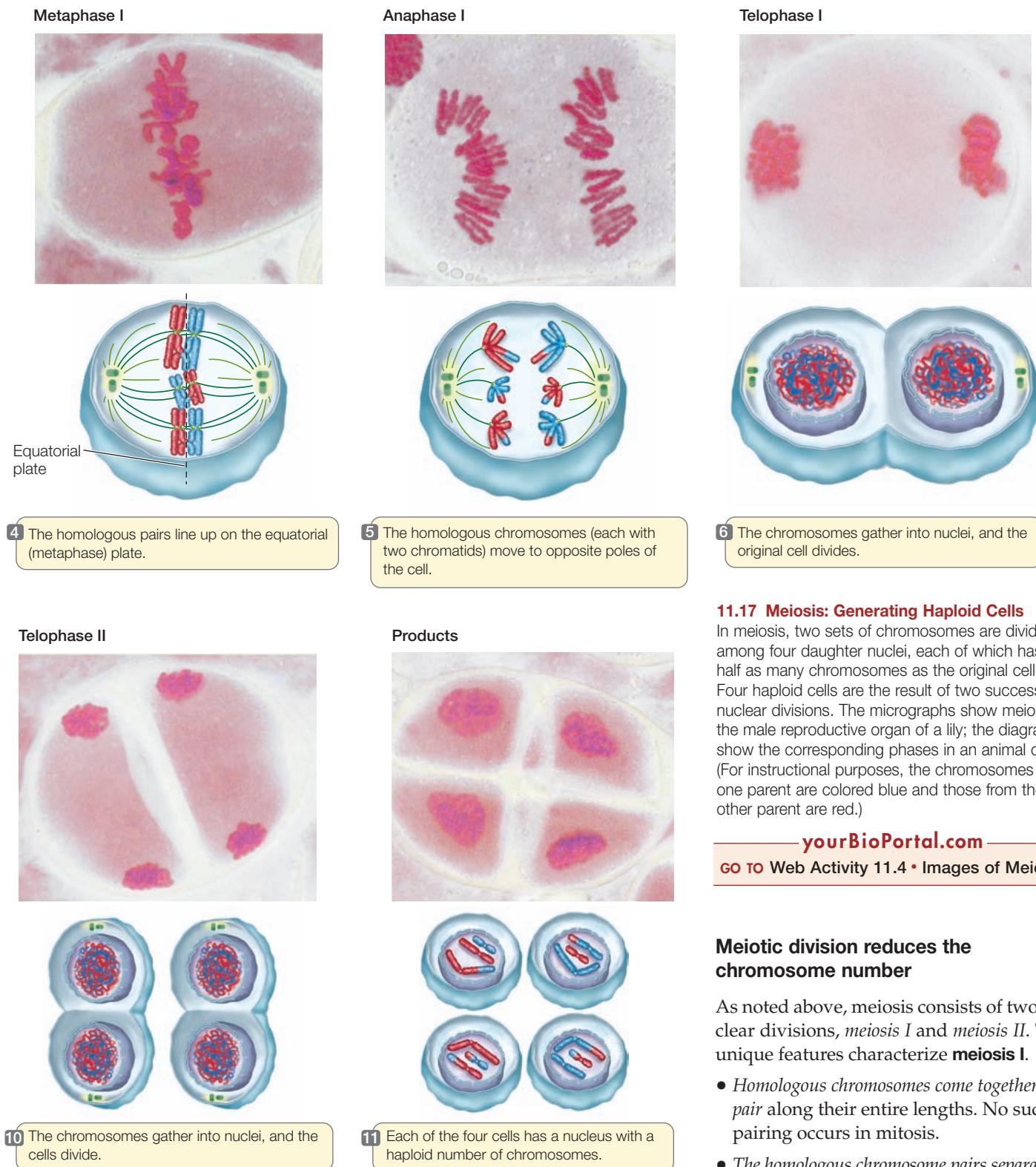
9 The chromatids finally separate, becoming chromosomes in their own right, and are pulled to opposite poles. Because of crossing over and independent assortment, each new cell will have a different genetic makeup.

11.5 What Happens during Meiosis?

In the last section we described the role and importance of meiosis in sexual reproduction. Now we will see how meiosis accomplishes the orderly and precise generation of haploid cells.

Meiosis consists of *two* nuclear divisions that reduce the number of chromosomes to the haploid number, in preparation

for sexual reproduction. Although the *nucleus divides twice* during meiosis, the *DNA is replicated only once*. Unlike the products of mitosis, the products of meiosis are genetically different from one another and from the parent cell. To understand the process of meiosis and its specific details, it is useful to keep in mind the overall functions of meiosis:



- To reduce the chromosome number from diploid to haploid
- To ensure that each of the haploid products has a complete set of chromosomes
- To generate genetic diversity among the products

The events of meiosis are illustrated in **Figure 11.17**. In this section, we discuss some of the key features that distinguish meiosis from mitosis.

11.17 Meiosis: Generating Haploid Cells

In meiosis, two sets of chromosomes are divided among four daughter nuclei, each of which has half as many chromosomes as the original cell. Four haploid cells are the result of two successive nuclear divisions. The micrographs show meiosis in the male reproductive organ of a lily; the diagrams show the corresponding phases in an animal cell. (For instructional purposes, the chromosomes from one parent are colored blue and those from the other parent are red.)

yourBioPortal.com

GO TO Web Activity 11.4 • Images of Meiosis

Meiotic division reduces the chromosome number

As noted above, meiosis consists of two nuclear divisions, *meiosis I* and *meiosis II*. Two unique features characterize **meiosis I**.

- *Homologous chromosomes come together to pair along their entire lengths. No such pairing occurs in mitosis.*
- *The homologous chromosome pairs separate, but the individual chromosomes, each consisting of two sister chromatids, remain intact. (The chromatids will separate during meiosis II.)*

Like mitosis, meiosis I is preceded by an interphase with an S phase, during which each chromosome is replicated. As a result, each chromosome consists of two sister chromatids, held together by cohesin proteins. At the end of meiosis I, two nuclei form, each with half of the original chromosomes (one member of each homologous pair). Since the centromeres did not

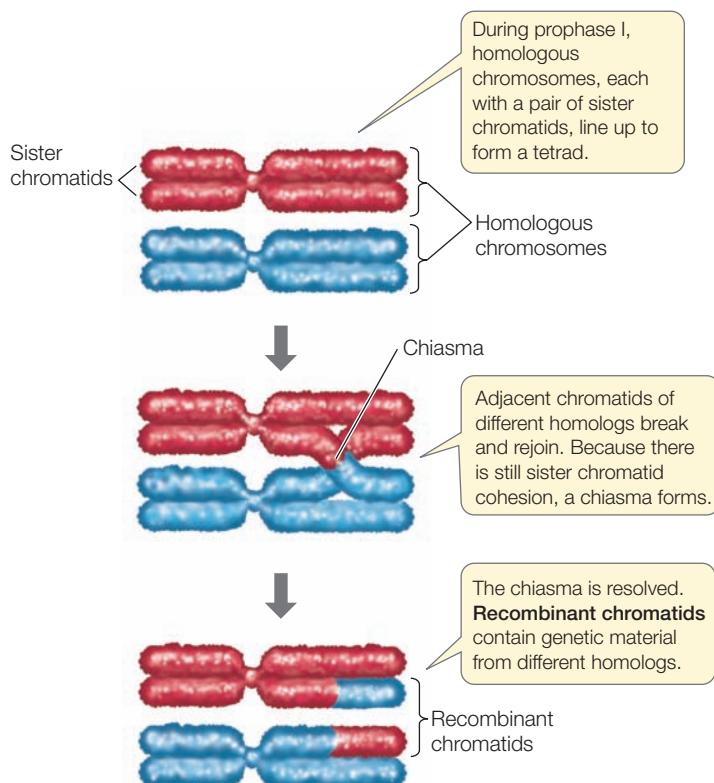
separate, these chromosomes are still double—composed of two sister chromatids. The sister chromatids are separated during **meiosis II**, which is *not* preceded by DNA replication. As a result, the products of meiosis I and II are four cells, each containing the haploid number of chromosomes. But these four cells are not genetically identical.

Chromatid exchanges during meiosis I generate genetic diversity

Meiosis I begins with a long prophase I (the first three panels of Figure 11.17), during which the chromosomes change markedly. The homologous chromosomes pair by adhering along their lengths in a process called **synapsis**. (This does not happen in mitosis.) This pairing process lasts from prophase I to the end of metaphase I. The four chromatids of each pair of homologous chromosomes form a **tetrad**, or *bivalent*. For example, in a human cell at the end of prophase I there are 23 tetrads, each consisting of four chromatids. The four chromatids come from the two partners in each homologous pair of chromosomes.

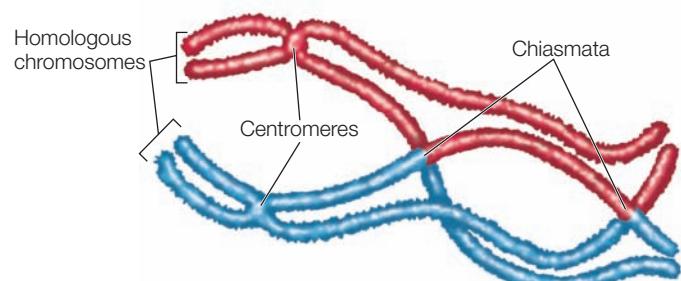
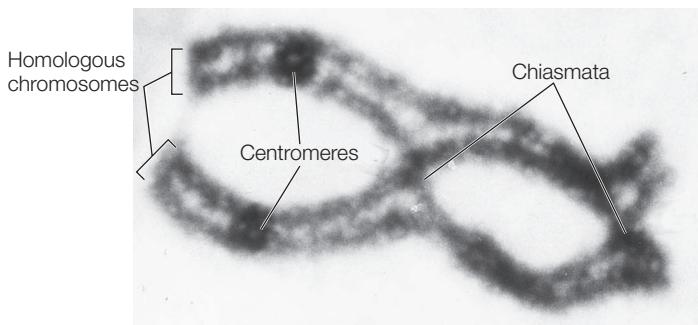
Throughout prophase I and metaphase I, the chromatin continues to coil and compact, so that the chromosomes appear ever thicker. At a certain point, the homologous chromosomes appear to repel each other, especially near the centromeres, but they remain held together by physical attachments mediated by cohesins. Later in prophase, regions having these attachments take on an X-shaped appearance (Figure 11.18) and are called **chiasmata** (singular *chiasma*, “cross”).

A chiasma reflects an *exchange of genetic material* between non-sister chromatids on homologous chromosomes—what geneticists call **crossing over** (Figure 11.19). The chromosomes usually



11.19 Crossing Over Forms Genetically Diverse Chromosomes

The exchange of genetic material by crossing over results in new combinations of genetic information on the recombinant chromosomes. The two different colors distinguish the chromosomes contributed by the male and female parents.

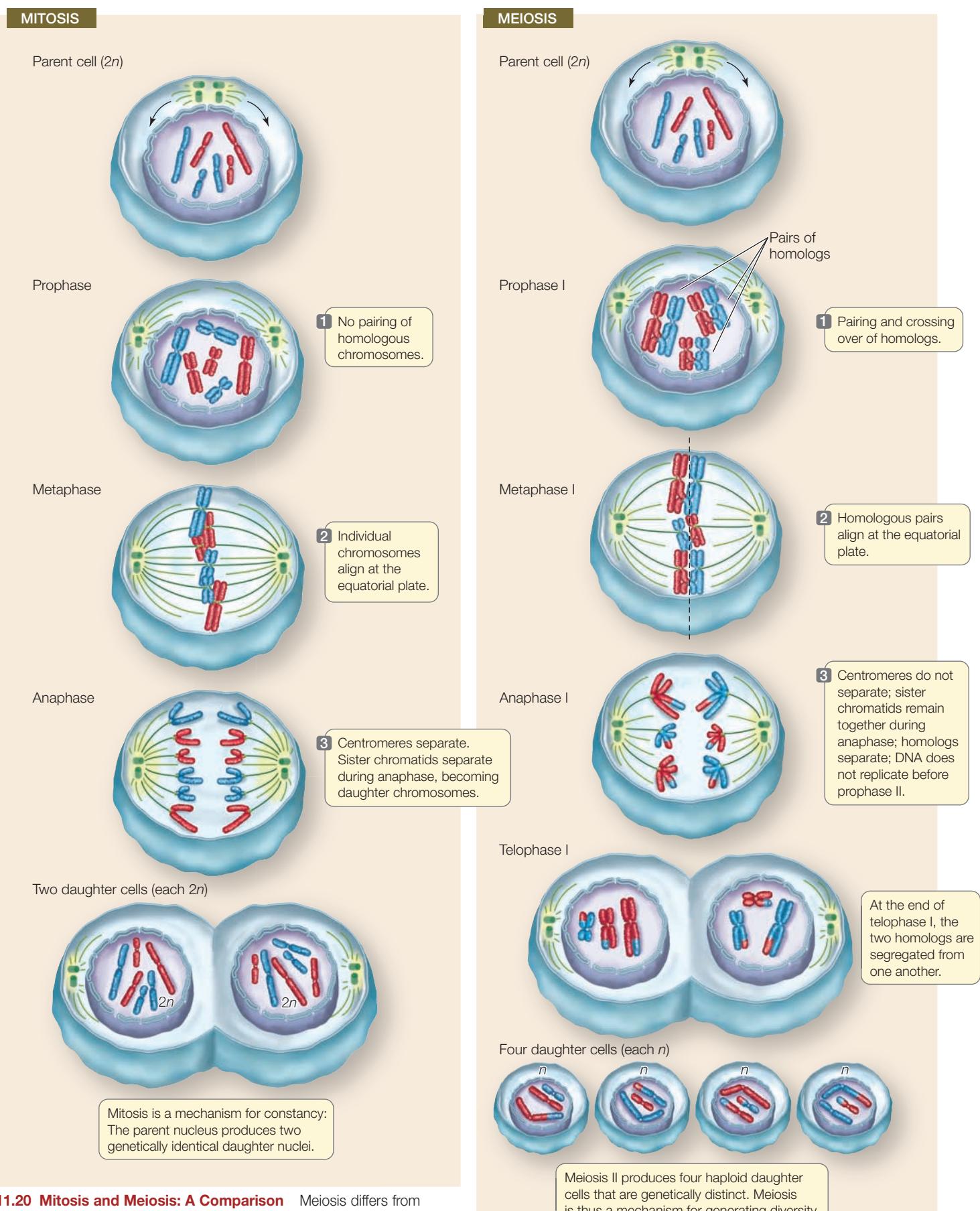


11.18 Chiasmata: Evidence of Genetic Exchange between Chromatids This micrograph shows a pair of homologous chromosomes, each with two chromatids, during prophase I of meiosis in a salamander. Two chiasmata are visible.

begin exchanging material shortly after synapsis begins, but chiasmata do not become visible until later, when the homologs are repelling each other. Crossing over results in **recombinant** chromatids, and it increases genetic variation among the products of meiosis by reshuffling genetic information among the homologous pairs. In Chapter 12 we explore further the genetic consequences of crossing over. Mitosis seldom takes more than an hour or two, but meiosis can take *much* longer. In human males, the cells in the testis that undergo meiosis take about a week for prophase I and about a month for the entire meiotic cycle. In females, prophase I begins long before a woman’s birth, during her early fetal development, and ends as much as decades later, during the monthly ovarian cycle.

During meiosis homologous chromosomes separate by independent assortment

A diploid organism has two sets of chromosomes ($2n$); one set derived from its male parent, and the other from its female parent. As the organism grows and develops, its cells undergo mitotic divisions. In mitosis, each chromosome behaves independently of its homolog, and its two chromatids are sent to opposite poles during anaphase. Each daughter nucleus ends up with $2n$ chromosomes. In meiosis, things are very different. Figure 11.20 compares the two processes.



11.20 Mitosis and Meiosis: A Comparison Meiosis differs from mitosis chiefly by the pairing of homologs and by the failure of the centromeres to separate at the end of metaphase I.

In meiosis I, chromosomes of maternal origin pair with their paternal homologs during synapsis. *This pairing does not occur in mitosis.* Segregation of the homologs during meiotic anaphase I (see steps 4–6 of Figure 11.17) ensures that each pole receives one member of each homologous pair. For example, at the end of meiosis I in humans, each daughter nucleus contains 23 of the original 46 chromosomes. In this way, the chromosome number is decreased from diploid to haploid. Furthermore, meiosis I guarantees that each daughter nucleus gets one full set of chromosomes.

Crossing over is one reason for the genetic diversity among the products of meiosis. The other source of diversity is independent assortment. It is a matter of chance which member of a homologous pair goes to which daughter cell at anaphase I. For example, imagine there are two homologous pairs of chromosomes in the diploid parent nucleus. A particular daughter nucleus could receive the paternal chromosome 1 and the maternal chromosome 2. Or it could get paternal 2 and maternal 1, or both maternal, or both paternal chromosomes. It all depends on the way in which the homologous pairs line up at metaphase I. This phenomenon is termed **independent assortment**.

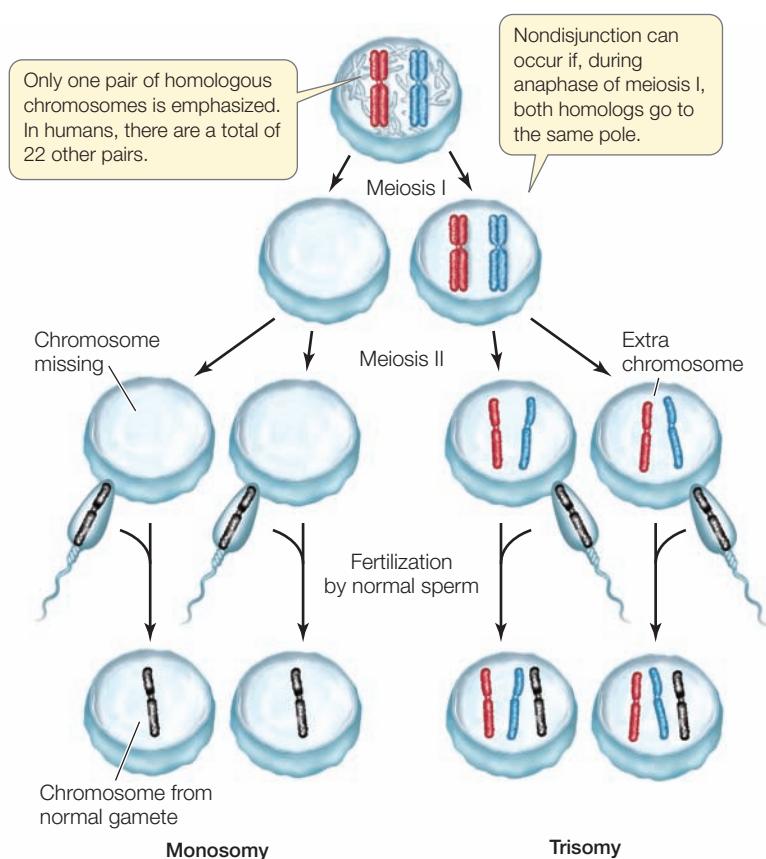
Note that of the four possible chromosome combinations just described, only two produce daughter nuclei with full complements of either maternal or paternal chromosome sets (apart from the material exchanged by crossing over). *The greater the number of chromosomes, the less probable that the original parental combinations will be reestablished, and the greater the potential for genetic diversity.* Most species of diploid organisms have more than two pairs of chromosomes. In humans, with 23 chromosome pairs, 2^{23} (8,388,608) different combinations can be produced just by the mechanism of independent assortment. Taking the extra genetic shuffling afforded by crossing over into account, the number of possible combinations is virtually infinite. Crossing over and independent assortment, along with the processes that result in mutations, provide the genetic diversity needed for evolution by natural selection.

We have seen how meiosis I is fundamentally different from mitosis. On the other hand, meiosis II is similar to mitosis, in that it involves the separation of chromatids into daughter nuclei (see steps 7–11 in Figure 11.17). The final products of meiosis I and meiosis II are four haploid daughter cells, each with one set (n) of chromosomes.

yourBioPortal.com
GO TO Animated Tutorial 11.2 • Meiosis

Meiotic errors lead to abnormal chromosome structures and numbers

In the complex process of meiosis, things occasionally go wrong. A pair of homologous chromosomes may fail to separate during meiosis I, or sister chromatids may fail to separate during meiosis II. Conversely, homologous chromosomes may fail to remain together during metaphase I, and then both may mi-



11.21 Nondisjunction Leads to Aneuploidy Nondisjunction occurs if homologous chromosomes fail to separate during meiosis I, or if chromatids fail to separate during meiosis II. The first case is shown here. The result is aneuploidy: one or more chromosomes are either lacking or present in excess. Generally, aneuploidy is lethal to the developing embryo.

grate to the same pole in anaphase I. This phenomenon is called **nondisjunction** and it results in the production of **aneuploid** cells. **Aneuploidy** is a condition in which one or more chromosomes are either lacking or present in excess (Figure 11.21).

There are many different causes of aneuploidy, but one of them may result from a breakdown in the cohesins that keep sister chromatids and tetrads joined together during prophase I (see Figure 11.17). These and other proteins ensure that when the chromosomes line up at the equatorial plate, one homolog will face one pole and the other homolog will face the other pole. If the cohesins break down at the wrong time, both homologs may go to one pole. If, for example, during the formation of a human egg, both members of the chromosome 21 pair go to the same pole during anaphase I, the resulting eggs will contain either two copies of chromosome 21 or none at all. If an egg with two of these chromosomes is fertilized by a normal sperm, the resulting zygote will have three copies of the chromosome: it will be **trisomic** for chromosome 21. A child with an extra chromosome 21 has the symptoms of *Down syndrome*: im-

paired intelligence; characteristic abnormalities of the hands, tongue, and eyelids; and an increased susceptibility to cardiac abnormalities and diseases such as leukemia. If an egg that did not receive chromosome 21 is fertilized by a normal sperm, the zygote will have only one copy: it will be **monosomic** for chromosome 21, and this is lethal.

Trisomies and the corresponding monosomies are surprisingly common in human zygotes, with 10–30 percent of all conceptions showing aneuploidy. But most of the embryos that develop from such zygotes do not survive to birth, and those that do often die before the age of 1 year (trisomies for chromosome 21 are the viable exception). At least one-fifth of all recognized pregnancies are spontaneously terminated (miscarried) during the first 2 months, largely because of trisomies and monosomies. The actual proportion of spontaneously terminated pregnancies is certainly higher, because the earliest ones often go unrecognized.

Other abnormal chromosomal events can also occur. In a process called **translocation**, a piece of a chromosome may break away and become attached to another chromosome. For example, a particular large part of one chromosome 21 may be translocated to another chromosome. Individuals who inherit this translocated piece along with two normal chromosomes 21 will have Down syndrome.

Polyploids have more than two complete sets of chromosomes

As mentioned in Section 11.4, mature organisms are often either diploid (for example, most animals) or haploid (for example, most fungi). Under some circumstances, triploid ($3n$), tetraploid ($4n$), or higher-order **polyploid** nuclei may form. Each of these *ploidy levels* represents an increase in the number of complete chromosome sets present. Organisms with complete extra sets of chromosomes may sometimes be produced by artificial breeding or by natural accidents. Polyploidy occurs naturally in some animals and many plants, and it has probably led to speciation (the evolution of a new species) in some cases.

A diploid nucleus can undergo normal meiosis because there are two sets of chromosomes to make up homologous pairs, which separate during anaphase I. Similarly, a tetraploid nucleus has an even number of each kind of chromosome, so each chromosome can pair with its homolog. However, a triploid nucleus cannot undergo normal meiosis because one-third of the chromosomes would lack partners. Polyploidy has implications for agriculture, particularly in the production of hybrid plants. For example, ploidy must be taken into account in wheat breeding because there are diploid, tetraploid, and hexaploid wheat varieties. Polyploidy can be a desirable trait in crops and ornamental plants because it often leads to more robust plants with larger flowers, fruits, and seeds. Triploidy can be useful in some circumstances. For example, rivers and lakes can be stocked with triploid trout, which are sterile and will not escape to reproduce in waters where they might upset the natural ecology.

11.5 RECAP

Meiosis produces four daughter cells in which the chromosome number is reduced from diploid to haploid. Because of the independent assortment of chromosomes and the crossing over of homologous chromatids, the four products of meiosis are not genetically identical. Meiotic errors, such as the failure of a homologous chromosome pair to separate, can lead to abnormal numbers of chromosomes.

- How do crossing over and independent assortment result in unique daughter nuclei? See p. 226 and Figure 11.19
- What are the differences between meiosis and mitosis? See pp. 224–228 and Figure 11.20
- What is aneuploidy, and how can it arise from nondisjunction during meiosis? See p. 228 and Figure 11.21

An essential role of cell division in complex eukaryotes is to replace cells that die. What happens to those cells?

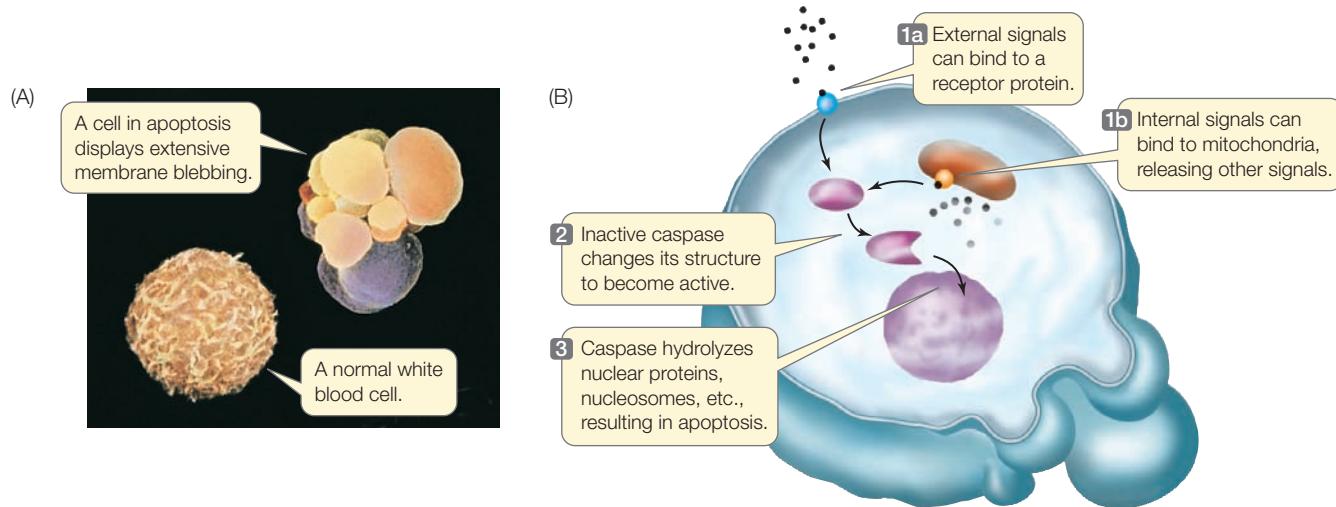
11.6 In a Living Organism, How Do Cells Die?

Cells die in one of two ways. The first type of cell death, **necrosis**, occurs when cells are damaged by mechanical means or toxins, or are starved of oxygen or nutrients. These cells usually swell up and burst, releasing their contents into the extracellular environment. This process often results in inflammation (see Section 42.2).

More typically, cell death is due to **apoptosis** (Greek, “falling apart”). Apoptosis is a programmed series of events that result in cell death. Why would a cell initiate apoptosis, which is essentially cell suicide? In animals, there are two possible reasons:

- *The cell is no longer needed by the organism.* For example, before birth, a human fetus has weblike hands, with connective tissue between the fingers. As development proceeds, this unneeded tissue disappears as its cells undergo apoptosis in response to specific signals.
- *The longer cells live, the more prone they are to genetic damage that could lead to cancer.* This is especially true of epithelial cells of the surface of an organism, which may be exposed to radiation or toxic substances. Such cells normally die after only days or weeks and are replaced by new cells.

The outward events of apoptosis are similar in many organisms. The cell becomes detached from its neighbors, cuts up its chromatin into nucleosome-sized pieces, and forms membranous lobes, or “blebs,” that break up into cell fragments (Figure 11.22A). In a remarkable example of the economy of nature, the surrounding living cells usually ingest the remains of the dead cell by phagocytosis. Neighboring cells digest the apop-



totic cell contents in their lysosomes and the digested components are recycled.

Apoptosis is also used by plant cells, in an important defense mechanism called the *hypersensitive response*. Plants can protect themselves from disease by undergoing apoptosis at the site of infection by a fungus or bacterium. With no living tissue to grow in, the invading organism is not able to spread to other parts of the plant. Because of their rigid cell wall, plant cells do not form blebs the way that animal cells do. Instead, they digest their own cell contents in the vacuole and then release the digested components into the vascular system.

Despite these differences between plant and animal cells, they share many of the signal transduction pathways that lead to apoptosis. Like the cell division cycle, programmed cell death is controlled by signals, which may come from inside or outside the cell (Figure 11.22B). Internal signals may be linked to the absence of mitosis or the recognition of damaged DNA. External signals (or a lack of them) can cause a receptor protein in the plasma membrane to change its shape, and in turn activate a signal transduction pathway. Both internal and external signals can lead to the activation of a class of enzymes called **caspases**. These enzymes are proteases that hydrolyze target molecules in a cascade of events. As a result, the cell dies as the caspases hydrolyze proteins of the nuclear envelope, nucleosomes, and plasma membrane.

11.22 Apoptosis: Programmed Cell Death (A) Many cells are programmed to “self-destruct” when they are no longer needed, or when they have lived long enough to accumulate a burden of DNA damage that might harm the organism. (B) Both external and internal signals stimulate caspases, the enzymes that break down specific cell constituents, resulting in apoptosis.

11.7 How Does Unregulated Cell Division Lead to Cancer?

Perhaps no malady affecting people in the industrialized world instills more fear than cancer, and most people realize that it involves an inappropriate increase in cell numbers. One in three Americans will have some form of cancer in their lifetimes, and at present, one in four will die of it. With 1.5 million new cases and half a million deaths in the United States annually, cancer ranks second only to heart disease as a killer.

Cancer cells differ from normal cells

Cancer cells differ from the normal cells from which they originate in two ways:

- Cancer cells lose control over cell division.
- Cancer cells can migrate to other locations in the body.

Most cells in the body divide only if they are exposed to extracellular signals such as growth factors. Cancer cells do not respond to these controls, and instead divide more or less continuously, ultimately forming **tumors** (large masses of cells). By the time a physician can feel a tumor or see one on an X-ray film or CAT scan, it already contains millions of cells. Tumors can be benign or malignant.

Benign tumors resemble the tissue they came from, grow slowly, and remain localized where they develop. For example, a lipoma is a benign tumor of fat cells that may arise in the armpit and remain there. Benign tumors are not cancers, but they must be removed if they impinge on an organ, obstructing its function.

11.6 RECAP

Cell death can occur either by necrosis or by apoptosis. Apoptosis is governed by precise molecular controls.

- What are some differences between apoptosis and necrosis? See p. 229
- In what situation is apoptosis necessary? See p. 229
- How is apoptosis regulated? See Figure 11.22

Malignant tumors do not look like their parent tissue at all. A flat, specialized epithelial cell in the lung wall may turn into a relatively featureless, round, malignant lung cancer cell (**Figure 11.23**). Malignant cells often have irregular structures, such as variable nucleus sizes and shapes. Recall the opening story of this chapter, in which cervical cancer was diagnosed by cell structure.

The second and most fearsome characteristic of cancer cells is their ability to invade surrounding tissues and spread to other parts of the body by traveling through the bloodstream or lymphatic ducts. When malignant cells become lodged in some distant part of the body they go on dividing and growing, establishing a tumor at that new site. This spreading, called **metastasis**, results in organ failures and makes the cancer very hard to treat.

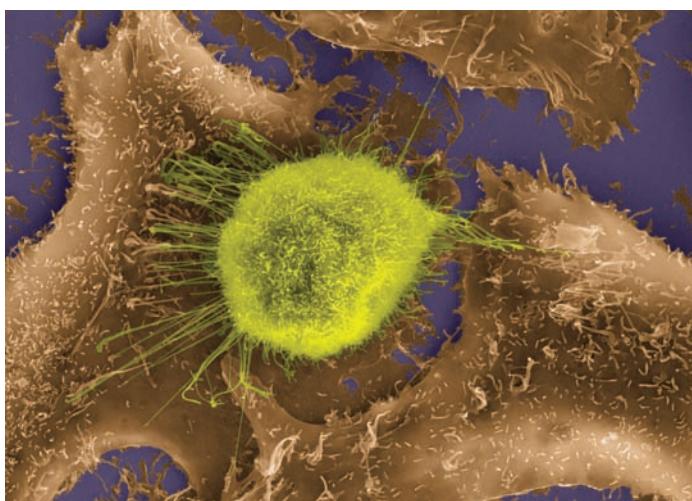
Cancer cells lose control over the cell cycle and apoptosis

Earlier in this chapter you learned about proteins that regulate the progress of a eukaryotic cell through the cell cycle:

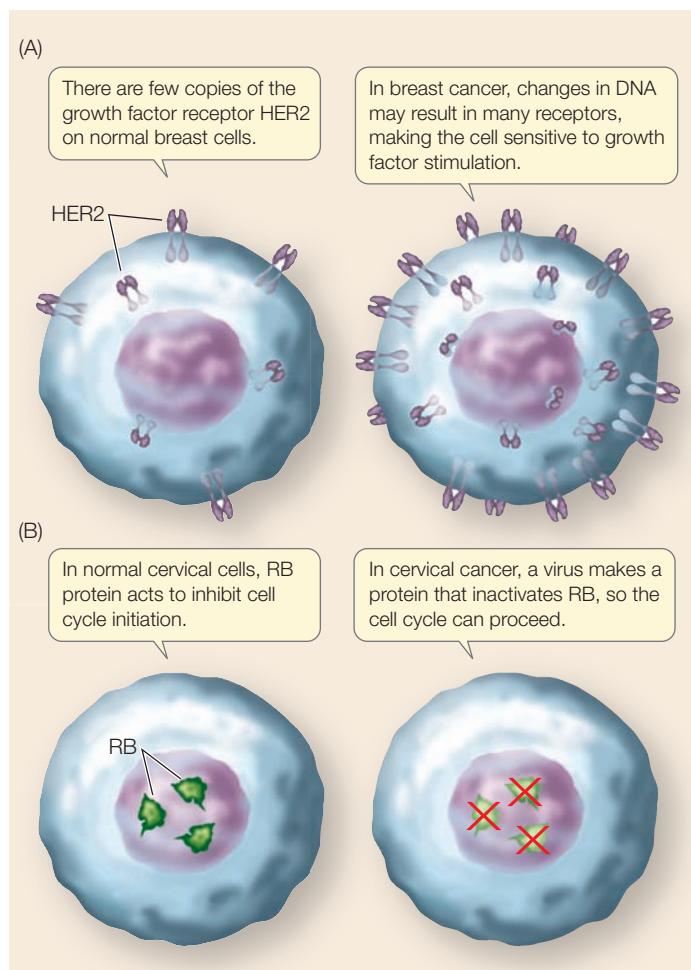
- Positive regulators such as growth factors stimulate the cell cycle: they are like “gas pedals.”
- Negative regulators such as RB inhibit the cell cycle: they are like “brakes.”

Just as driving a car requires stepping on the gas pedal *and* releasing the brakes, a cell will go through a division cycle only if the positive regulators are active and the negative regulators are inactive.

In most cells, the two regulatory systems ensure that cells divide only when needed. In cancer cells, these two processes are abnormal.



11.23 A Cancer Cell with its Normal Neighbors This lung cancer cell (yellow-green) is quite different from the normal lung cells surrounding it. The cancer cell can divide more rapidly than its normal counterparts, and it can spread to other organs. This form of small-cell cancer is lethal, with a 5-year survival rate of 10 percent. Most cases are caused by tobacco smoking.



11.24 Molecular Changes in Cancer Cells In cancer, oncogene proteins become active (A) and tumor suppressor proteins become inactive (B).

- **Oncogene** proteins are positive regulators in cancer cells. They are derived from normal positive regulators that have become mutated to be overly active or that are present in excess, and they stimulate the cancer cells to divide more often. Oncogene products could be growth factors, their receptors, or other components in the signal transduction pathway. An example of an oncogene protein is the growth factor receptor in a breast cancer cell (**Figure 11.24A**). Normal breast cells have relatively low numbers of the growth factor receptor HER2. So when this growth factor is made, it doesn't find many breast cell receptors with which to bind and initiate cell division. In about 25 percent of breast cancers, a DNA change results in the increased production of the HER2 receptor. This results in positive stimulation of the cell cycle, and a rapid proliferation of cells with the altered DNA.

- **Tumor suppressors** are negative regulators in both cancer and normal cells, but in cancer cells they are inactive. An example is the RB protein that acts at R (the restriction point) in G1 (see Figure 11.6). When RB is active the cell cycle does not proceed, but it is inactive in cancer cells, allowing the

cell cycle to occur. Some viral proteins can inactivate tumor suppressors. For example, in the opening story of this chapter we saw how HPV infects cells of the cervix and produces a protein called E7. E7 binds to the RB protein and prevents it from inhibiting the cell cycle (Figure 11.24B).

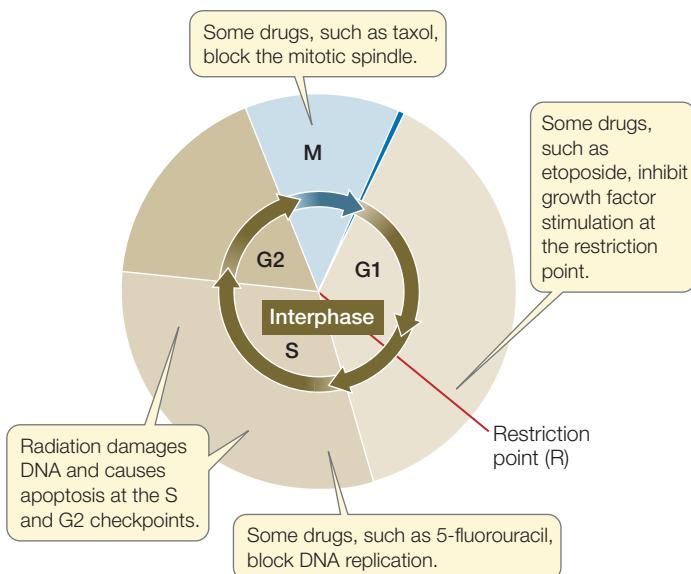
The discovery of apoptosis and its importance (see Section 11.6) has changed the way biologists think about cancer. In a population of organisms, the net increase in the number of individuals over time (the growth rate) is a function of the individuals added (the birth rate) and lost (the death rate). Cell populations behave the same way:

$$\text{growth rate of cell population} = \frac{\text{rate of cell division ("births")}}{\text{rate of apoptosis ("deaths")}}$$

Cancer cells may lose the ability to respond to positive regulators of apoptosis (see Figure 11.22). This lowers the cellular “death rate” so that the overall cell population grows rapidly.

Cancer treatments target the cell cycle

The most successful and widely used treatment for cancer is surgery. While physically removing a tumor is optimal, it is often difficult for a surgeon to get all of the tumor cells. (A tumor about 1 cm in size already has a billion cells!) Tumors are generally embedded in normal tissues. Added to this is the probability that cells of the tumor may have broken off and



11.25 Cancer Treatment and the Cell Cycle To prevent cancer cells from dividing, physicians use combinations of therapies that attack the cell cycle at different points.

spread to other organs. This makes it unlikely that localized surgery will be curative. So other approaches are taken to treat or cure cancer, and these generally target the cell cycle (Figure 11.25).

An example of a cancer drug that targets the cell cycle is *5-fluorouracil*, which blocks the synthesis of thymine, one of the four bases in DNA. The drug *taxol* prevents the functioning of microtubules in the mitotic spindle. Both drugs inhibit the cell cycle, and apoptosis causes tumor shrinkage. More dramatic is radiation treatment, in which a beam of high-energy radiation is focused on the tumor. DNA damage is extensive, and the cell cycle checkpoint for DNA repair is overwhelmed. As a result, the cell undergoes apoptosis. A major problem with these treatments is that they target normal cells as well as the tumor cells. These treatments are toxic to tissues with large populations of normal dividing cells such as those in the intestine, skin and bone marrow (producing blood cells).

A major effort in cancer research is to find treatments that target only cancer cells. A promising recent example is *Herceptin*, which targets the HER2 growth factor receptor that is expressed at high levels on the surfaces of some breast cancer cells (see Figure 11.24A). Herceptin binds specifically to the HER2 receptor but does not stimulate it. This prevents the natural growth factor from binding, and so the cells are not stimulated to divide. As a result, the tumor shrinks because the apoptosis rate remains the same. More such treatments are on the way.

11.7 RECAP

Cancer cells differ from normal cells in terms of their rapid cell division and their ability to spread (metastasis). Many proteins regulate the cell cycle, either positively or negatively. In cancer, one or another of these proteins is altered in some way, making its activity abnormal. Radiation and many cancer drugs target proteins involved in the cell cycle.

- How are oncogene proteins and tumor suppressor proteins involved in cell cycle control in normal and cancer cells? Review p. 231 and Figure 11.24
- How does cancer treatment target the cell cycle? Review p. 232 and Figure 11.25

We have now looked at the cell cycle and at cell division by binary fission, mitosis, and meiosis. We have described the normal cell cycle and how it is upset in cancer. We have seen how meiosis produces haploid cells in sexual life cycles. In the coming chapters we examine heredity, genes, and DNA. In Chapter 12 we see how Gregor Mendel studied heredity in the nineteenth century and how the enormous power of his discoveries founded the science of genetics, and changed forever the science of biology.

CHAPTER SUMMARY

11.1 How Do Prokaryotic and Eukaryotic Cells Divide?

- Cell division is necessary for the reproduction, growth, and repair of organisms.
- Cell division must be initiated by a reproductive signal. Before a cell can divide, the genetic material (DNA) must be **replicated** and **segregated** to separate portions of the cell. **Cytokinesis** then divides the cytoplasm into two cells.
- In prokaryotes, most cellular DNA is a single molecule, usually in the form of a circular **chromosome**. Prokaryotes reproduce by **binary fission**. [Review Figure 11.2](#)
- In eukaryotes, cells divide by either **mitosis** or **meiosis**. Eukaryotic cell division follows the same general pattern as binary fission, but with significant differences. For example, a eukaryotic cell has a distinct nucleus that must be replicated prior to separating the two daughter cells.
- Cells that produce gametes undergo a special kind of nuclear division called meiosis; the four daughter cells produced by meiosis are not genetically identical.

11.2 How Is Eukaryotic Cell Division Controlled?

- The eukaryotic cell cycle has two main phases: **interphase**, during which cells are not dividing and the DNA is replicating, and **mitosis** or **M phase**, when the cells are dividing.
- During most of the eukaryotic cell cycle, the cell is in interphase, which is divided into three subphases: **S**, **G₁**, and **G₂**. DNA is replicated during the S phase. Mitosis (M phase) and cytokinesis follow. [Review Figure 11.3](#)
- Cyclin–Cdk complexes** regulate the passage of cells through checkpoints in the cell cycle. The suppressor protein **RB** inhibits the cell cycle. The G₁–S cyclin–Cdk functions by inactivating RB and allows the cell cycle to progress beyond the **restriction point**. [Review Figures 11.5 and 11.6](#)
- External controls such as **growth factors** can also stimulate the cell to begin a division cycle.

11.3 What Happens during Mitosis?

SEE ANIMATED TUTORIAL 11.1

- In mitosis, a single nucleus gives rise to two nuclei that are genetically identical to each other and to the parent nucleus.
- DNA is wrapped around proteins called **histones**, forming beadlike units called **nucleosomes**. A eukaryotic chromosome contains strings of nucleosomes bound to proteins in a complex called **chromatin**. [Review Figure 11.9](#)
- At mitosis, the replicated chromosomes, called **sister chromatids**, are held together at the **centromere**. Each chromatid consists of one double-stranded DNA molecule. [Review Figure 11.10, WEB ACTIVITY 11.1](#)
- Mitosis can be divided into several phases called **prophase**, **prometaphase**, **metaphase**, **anaphase**, and **telophase**.
- During mitosis sister chromatids, attached by **cohesin**, line up at the equatorial plate and attach to the **spindle**. The chromatids separate (becoming **daughter chromosomes**) and migrate to opposite ends of the cell. [Review Figure 11.11, WEB ACTIVITY 11.2](#)

- Nuclear division is usually followed by cytokinesis. Animal cell cytoplasms divide via a contractile ring made up of actin microfilaments. In plant cells, cytokinesis is accomplished by vesicles that fuse to form a cell plate. [Review Figure 11.13](#)

11.4 What Role Does Cell Division Play in a Sexual Life Cycle?

- Asexual reproduction** produces clones, new organisms that are genetically identical to the parent. Any genetic variation is the result of mutations.
- In **sexual reproduction**, two **haploid** gametes—one from each parent—unite in **fertilization** to form a genetically unique, **diploid zygote**. There are many different sexual life cycles that can be **haplontic**, **diplontic**, or involve **alternation of generations**. [Review Figure 11.15, WEB ACTIVITY 11.3](#)
- In sexually reproducing organisms, certain cells in the adult undergo meiosis, a process by which a diploid cell produces haploid gametes.
- Each gamete contains one of each pair of **homologous chromosomes** from the parent.
- The numbers, shapes, and sizes of the chromosomes constitute the **karyotype** of an organism.

11.5 What Happens during Meiosis?

SEE ANIMATED TUTORIAL 11.2

- Meiosis consists of two nuclear divisions, **meiosis I** and **meiosis II**, that collectively reduce the chromosome number from diploid to haploid. It ensures that each haploid cell contains one member of each chromosome pair, and results in four genetically diverse haploid cells, usually gametes. [Review Figure 11.17, WEB ACTIVITY 11.4](#)
- In meiosis I, entire chromosomes, each with two chromatids, migrate to the poles. In meiosis II, the sister chromatids separate.
- During prophase I, homologous chromosomes undergo **synapsis** to form pairs in a **tetrad**. Chromatids can form junctions called **chiasmata** and genetic material may be exchanged between the two homologs by **crossing over**. [Review Figure 11.18](#)
- Both crossing over during prophase I and **independent assortment** of the homologs as they separate during anaphase I ensure that the gametes are genetically diverse.
- In **nondisjunction**, two members of a homologous pair of chromosomes go to the same pole during meiosis I, or two chromatids go to the same pole during meiosis II. This leads to one gamete having an extra chromosome and another lacking that chromosome. [Review Figure 11.21](#)
- The union between a gamete with an abnormal chromosome number and a normal haploid gamete results in **aneuploidy**. Such genetic abnormalities are harmful or lethal to the organism.

11.6 In a Living Organism, How Do Cells Die?

- A cell may die by **necrosis**, or it may self-destruct by **apoptosis**, a genetically programmed series of events that includes the fragmentation of its nuclear DNA.

- Apoptosis is regulated by external and internal signals. These signals result in activation of a class of enzymes called **caspases** that hydrolyze proteins in the cell. **Review Figure 11.22**

11.7 How Does Unregulated Cell Division Lead to Cancer?

- Cancer cells divide more rapidly than normal cells and can be **metastatic**, spreading to distant organs in the body.

- Cancer can result from changes in either of two types of proteins that regulate the cell cycle. **Oncogene** proteins stimulate cell division and are activated in cancer. **Tumor suppressor** proteins normally inhibit the cell cycle but in cancer they are inactive. **Review Figure 11.24**

- Cancer treatment often targets the cell cycle in tumor cells. **Review Figure 11.25**

SELF-QUIZ

- Which statement about eukaryotic chromosomes is *not* true?
 - They sometimes consist of two chromatids.
 - They sometimes consist only of a single chromatid.
 - They normally possess a single centromere.
 - They consist only of proteins.
 - During metaphase they are visible under the light microscope.
- Nucleosomes
 - are made of chromosomes.
 - consist entirely of DNA.
 - consist of DNA wound around a histone core.
 - are present only during mitosis.
 - are present only during prophase.
- Which statement about the cell cycle is *not* true?
 - It consists of interphase, mitosis, and cytokinesis.
 - The cell's DNA replicates during G1.
 - A cell can remain in G1 for weeks or much longer.
 - DNA is not replicated during G2.
 - Cells enter the cell cycle as a result of internal or external signals.
- Which statement about mitosis is *not* true?
 - A single nucleus gives rise to two identical daughter nuclei.
 - The daughter nuclei are genetically identical to the parent nucleus.
 - The centromeres separate at the onset of anaphase.
 - Homologous chromosomes synapse in prophase.
 - The centrosomes organize the microtubules of the spindle fibers.
- Which statement about cytokinesis is true?
 - In animals, a cell plate forms.
 - In plants, it is initiated by furrowing of the membrane.
 - It follows mitosis.
 - In plant cells, actin and myosin play an important part.
 - It is the division of the nucleus.
- Apoptosis
 - occurs in all cells.
 - involves the formation of the plasma membrane.
 - does not occur in an embryo.
 - is a series of programmed events resulting in cell death.
 - is the same as necrosis.
- In meiosis,
 - meiosis II reduces the chromosome number from diploid to haploid.
 - DNA replicates between meiosis I and meiosis II.
 - the chromatids that make up a chromosome in meiosis II are identical.
 - each chromosome in prophase I consists of four chromatids.
 - homologous chromosomes separate from one another in anaphase I.
- In meiosis,
 - a single nucleus gives rise to two daughter nuclei.
 - the daughter nuclei are genetically identical to the parent nucleus.
 - the centromeres separate at the onset of anaphase I.
 - homologous chromosomes synapse in prophase I.
 - no spindle forms.
- An animal has a diploid chromosome number of 12. An egg cell of that animal has 5 chromosomes. The most probable explanation is
 - normal mitosis.
 - normal meiosis.
 - nondisjunction in meiosis I.
 - nondisjunction in meiosis I or II.
 - nondisjunction in mitosis.
- The number of daughter chromosomes in a human cell (diploid number 46) in anaphase II of meiosis is
 - 2.
 - 23.
 - 46.
 - 69.
 - 92.

FOR DISCUSSION

- Compare the roles of cohesins in mitosis, meiosis I, and meiosis II.
- Compare and contrast cell division in animals and plants.
- Contrast mitotic prophase and prophase I of meiosis. Contrast mitotic anaphase and anaphase I of meiosis.
- Compare the sequence of events in the mitotic cell cycle with the sequence of events in apoptosis.
- Cancer-fighting drugs are rarely used alone. Usually, there are several drugs given in combination that target different stages of the cell cycle. Why might this be a better approach than single drugs?

ADDITIONAL INVESTIGATION

1. Suggest two ways in which one might use a microscope to determine the relative durations of the various phases of mitosis.
2. Studying the events and controls of the cell cycle is much easier if the cells under investigation are synchronous; that is, if they are all in the same stage of the cell cycle. This can be accomplished with various chemicals. But some populations of cells are naturally synchronous. The anther (male

sex organ) of a lily plant contains cells that become pollen grains (male gametes). As anthers develop in the flower, their lengths correlate precisely with the stage of the meiotic cycle in those cells. These stages each take many days, so an anther that is 1.5 millimeters long, for example, contains cells in early prophase I. How would you use lily anthers to investigate the roles of cyclins and Cdk's in the meiotic cell cycle?

12

Inheritance, Genes, and Chromosomes

Genetic piracy

The Nazis were infamous for plundering the art collections of Europe during World War II, but why did they also spirit away a collection of seeds? The answer lies in the power and promise offered by the scientific understanding of genetics, the science of *heredity*.

A key step in the rise of human civilizations was the development of agriculture—the cultivation of plants and animals for food and other human needs. Some 10,000 years ago, early farmers began preferentially cultivating plants with certain traits (e.g., that survived drought better). Over time, the cultivated varieties (*cultivars*) became quite different from their wild relatives, an example of evolution by selection. In this case, it was not the result of

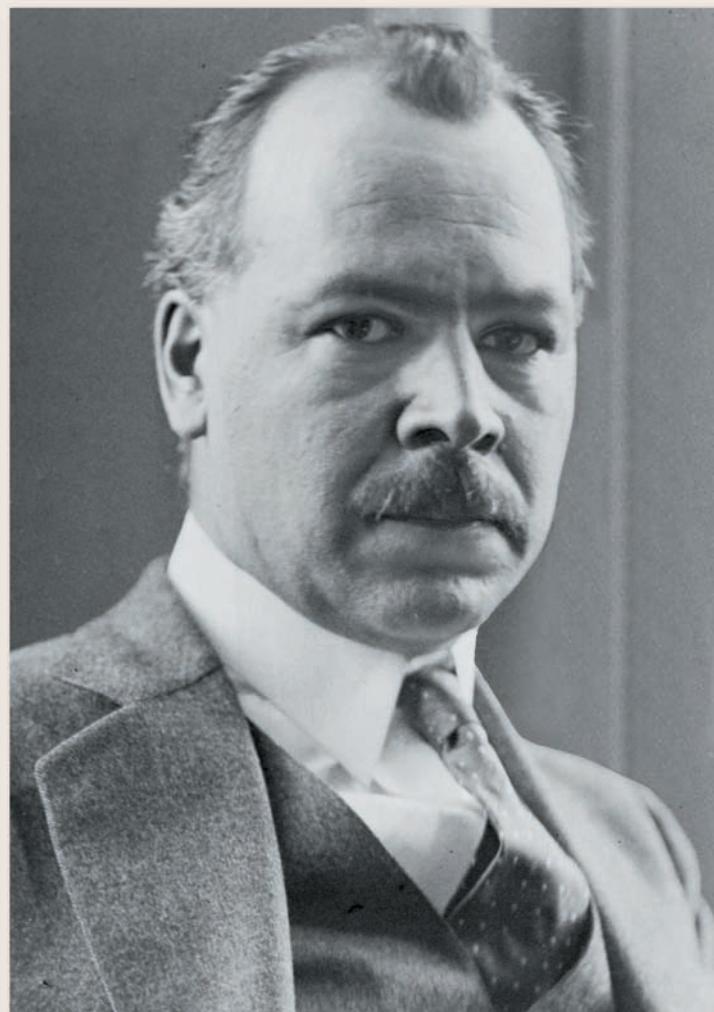
natural selection (see Chapter 21), but of artificial selection by the practices of ancient farmers.

Early in the twentieth century, Russian scientist Nicolai Vavilov began systematically collecting seeds from thousands of cultivars and their wild relatives. He convinced Lenin, the leader of the new Communist regime, that his seed collection would be useful in breeding crops that would be more productive in the difficult Russian climate. Lenin put Vavilov in charge of a large research institute. But when Lenin died in 1929, his successor, Josef Stalin, had little interest in science.

A politically ambitious student of Vavilov's, Trofim Lysenko, proposed to Stalin that favorable characteristics in plants could be rendered heritable by manipulating the parent plant's *phenotype* (physical state). This idea was at odds with what scientists knew about heredity and evolution, but it appealed to Stalin's political ideology. Stalin put Lysenko in charge of Vavilov's institute and sent Vavilov to a prison camp, where he died in 1943. Vavilov's unique seed collection—a *gene bank*—was ignored.

Meanwhile, in Germany, the Nazi leader Heinrich Himmler learned of the collection and was convinced that Vavilov's seeds could be a valuable key to providing better crops for the expanding German empire. Himmler put Heinz Brücher, a young SS officer with a doctorate in botany, in charge of obtaining the seeds. When the German army invaded Russia, Brücher's team removed thousands of seeds to a castle in Austria that already housed a seed collection Brücher had brought from Tibet.

Brücher's aim was to cross-breed plants from Tibet with plants from Russia to develop new crops that would grow well at high elevations and in cold climates; these plans came to a halt with the end of World War II. However,



Genetics Pioneer Collecting thousands of crop plant varieties from all over the world, Nikolai Vavilov laid the foundations for theories about the genetic origins of modern crops.



Hardy Grain Early geneticists hoped to increase food production by breeding crop varieties adapted to harsh climates (such as those in Tibet) with varieties with other desirable traits.

Brücher ignored a superior's order to blow up the castle, thus preserving most of Vavilov's seed bank. The collection was returned to Russia, where it continued to be used in breeding programs.

The ideas of Vavilov and the breeding plans of Brücher depended on the principles of *genetics*, a science born in an Austrian monastery in the 1860s, where Gregor Mendel performed—and, importantly, correctly interpreted—experiments on pea plants. It was almost 50 years before the scientific community recognized the significance of Mendel's work, but once that recognition was achieved, science and medicine sprang forward at a rapid pace.

IN THIS CHAPTER we will discuss how the units of inheritance—genes—are transmitted from generation to generation. We will show that many of the rules that govern inheritance can be explained by the behavior of chromosomes during meiosis. We will describe the interactions of genes with one another and with the environment, and we will see how the specific positions of genes on chromosomes affect diversity.

CHAPTER OUTLINE

- 12.1 What Are the Mendelian Laws of Inheritance?
- 12.2 How Do Alleles Interact?
- 12.3 How Do Genes Interact?
- 12.4 What Is the Relationship between Genes and Chromosomes?
- 12.5 What Are the Effects of Genes Outside the Nucleus?
- 12.6 How Do Prokaryotes Transmit Genes?

12.1 What Are the Mendelian Laws of Inheritance?

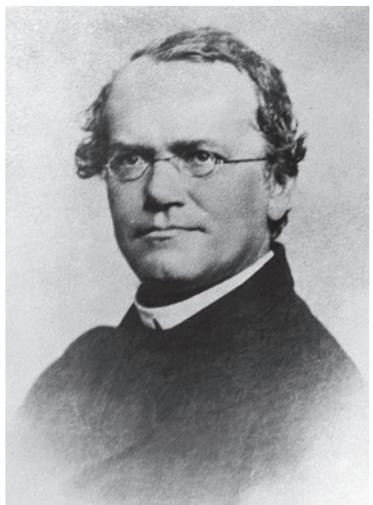
Much of the early study of biological inheritance was done with plants and animals of economic importance. Records show that people were deliberately cross-breeding date palm trees and horses as early as 5,000 years ago. By the early nineteenth century plant breeding was widespread, especially for ornamental flowers such as tulips. Plant breeders of that time were operating under two key assumptions about how inheritance worked. Only one of those assumptions turned out to be supported by experimental evidence.

- *Each parent contributes equally to offspring* (supported by experiments). In the 1770s, the German botanist Josef Gottlieb Kölreuter studied the offspring of **reciprocal crosses**, in which plants were crossed (mated with each other) in both directions. For example, in one cross, plants with white flowers were used as males to pollinate related plants with red flowers. In the complementary crosses, the red-flowered plants were used as males in crosses with the white flowered plants. In Kölreuter's studies, such reciprocal crosses always gave identical results, showing that both parents contributed equally to the offspring.
- *Hereditary determinants blend in offspring* (not supported by experiments). Kölreuter and others proposed that there were hereditary determinants in the egg and sperm cells. When these determinants came together in a single cell after mating, they were believed to blend together. If a plant with one form of a character (say, red flowers) was crossed with a plant with a different form of that character (blue flowers), the offspring would have a blended combination of the two parents' characteristics (purple flowers). According to the blending theory, once heritable elements were combined, they could not be separated again (like inks of different colors mixed together). The red and blue hereditary determinants were thought to be forever blended into the new purple one.

In his experiments in the 1860s, Gregor Mendel confirmed the first of these two assumptions but refuted the second.

Mendel brought new methods to experiments on inheritance

Gregor Mendel was an Austrian monk, not an academic scientist (**Figure 12.1**). He was well qualified, however, to under-



12.1 Gregor Mendel and His Garden The Austrian monk Gregor Mendel (left) did his groundbreaking genetics experiments in a garden at the monastery at Brno, in what is now the Czech Republic.



take scientific investigations. In 1850 he failed an examination for a teaching certificate in natural science, so he undertook intensive studies in physics, chemistry, mathematics, and various aspects of biology at the University of Vienna. His studies in physics and mathematics under the famous physicist Christian Doppler strongly influenced his use of experimental and quantitative methods in his studies of heredity, and it was those quantitative experiments that were key to his successful deductions.

Over the seven years he spent working out some principles of inheritance in plants, Mendel made crosses with hundreds of plants and noted the resulting characteristics of 24,034 progeny. Analysis of his meticulously gathered data suggested to him a new theory of how inheritance might work. He presented this theory in a public lecture in 1865 and a detailed written publication in 1866. Mendel's paper appeared in a journal that was received by 120 libraries, and he sent reprinted copies (of which he had obtained 40) to several distinguished scholars, including Charles Darwin. However, his theory was not readily accepted. In fact, it was mostly ignored.

One reason Mendel's paper received so little attention was that most prominent biologists of his time were not in the habit of thinking in mathematical terms, even the simple terms Mendel used. Even Charles Darwin, whose theory of evolution by natural selection was predicated on heritable variations among individuals, failed to understand the significance of

Mendel's findings. In fact, Darwin performed breeding experiments on snapdragons that were similar to Mendel's work with peas. Although Darwin's data were similar to Mendel's, he failed to question the assumption that parental contributions blend in the offspring.

By 1900, the events of meiosis had been observed and described, and Mendel's discoveries burst into sudden prominence as a result of independent experiments by three plant geneticists: Hugo DeVries, Carl Correns, and Erich von Tschermark. Each carried out crossing experiments, each published his principal findings in 1900, and each cited Mendel's 1866 paper. These three men realized that chromosomes and meiosis provided a physical explanation for the theory that Mendel had proposed to explain the data from his crosses.

That Mendel was able to achieve his remarkable insights before the discovery of genes and meiosis was largely due to his experimental methods. His work is a definitive example of extensive preparation, meticulous execution, and imaginative yet logical interpretation. He was also fortunate in his choice of experimental subjects. Let's take a closer look at these experiments and the conclusions and hypotheses that emerged.

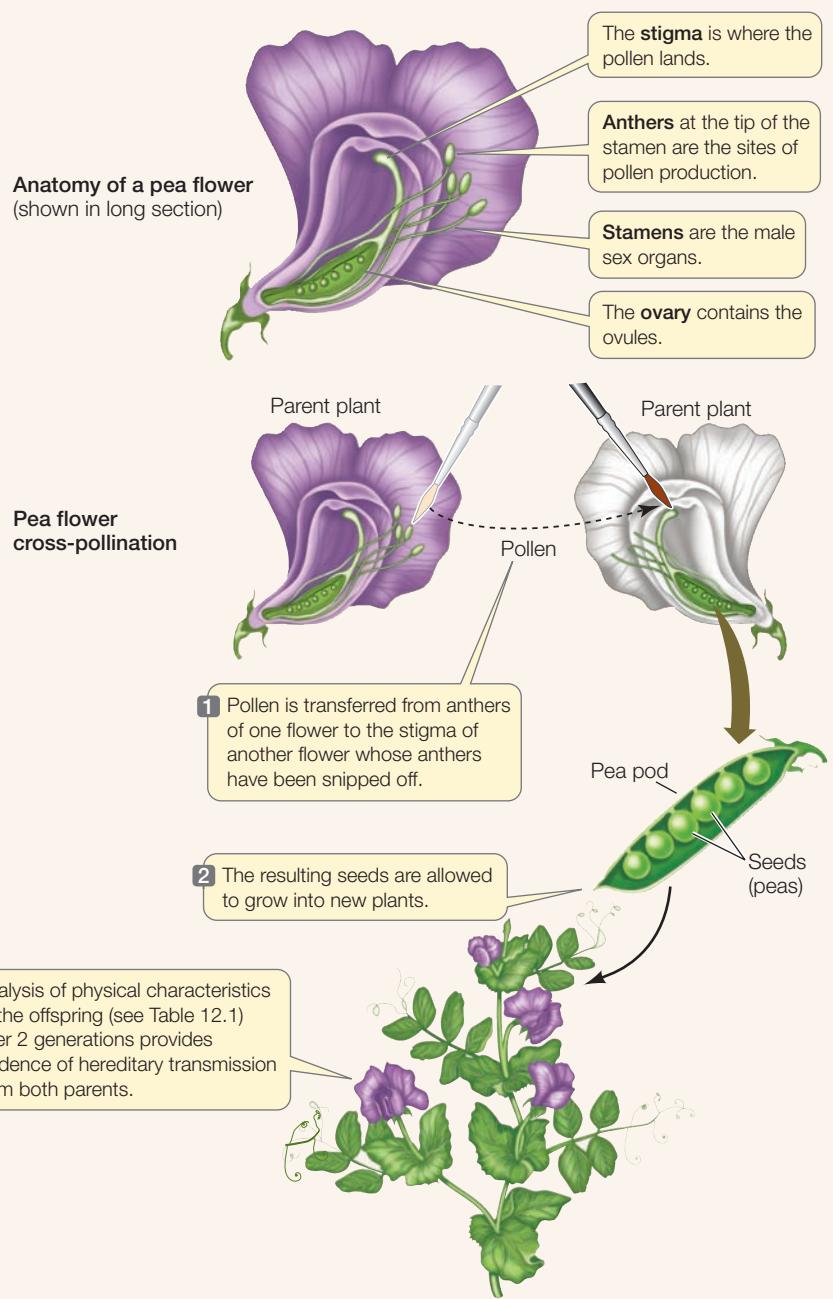
Mendel devised a careful research plan

Mendel chose to study the common garden pea because of its ease of cultivation, the feasibility of controlled pollination, and

TOOLS FOR INVESTIGATING LIFE

12.2 A Controlled Cross between Two Plants

Plants were widely used in early genetic studies because it is easy to control which individuals mate with which. Mendel used the garden pea (*Pisum sativum*) in many of his experiments.



the availability of varieties with contrasting traits. He controlled pollination, and thus fertilization, of his parent plants by manually moving pollen from one plant to another (Figure 12.2). Thus he knew the parentage of the offspring in his experiments. The pea plants Mendel studied produce male and female sex organs and gametes—sex cells such as eggs and sperm—in the

same flower. If untouched, they naturally self-pollinate—that is, the female organ of each flower receives pollen from the male organs of the same flower. Mendel made use of this natural phenomenon in some of his experiments.

Mendel began by examining different varieties of peas in a search for heritable characters and traits suitable for study:

- A **character** is an observable physical feature, such as flower color.
- A **trait** is a particular form of a character, such as purple flowers or white flowers.
- A **heritable trait** is one that is passed from parent to offspring.

Mendel looked for characters with well-defined, contrasting alternative traits, such as purple flowers versus white flowers. Furthermore, these traits had to be **true-breeding**, meaning that the observed trait was the only form present for many generations. In other words, if they were true-breeding, peas with white flowers would give rise only to progeny with white flowers when self-pollinated or crossed with one another for repeated generations. Similarly, tall plants bred with other tall plants would produce only tall progeny.

Mendel isolated each of his true-breeding strains by repeated inbreeding (done by crossing sibling plants that were seemingly identical or by allowing individuals to self-pollinate) and selection. In most of his work, Mendel concentrated on the seven pairs of contrasting traits shown in Table 12.1 (left side). His use of true-breeding strains for experimental crosses was an essential feature of his work.

Mendel then performed his crosses in the following manner:

- He removed the anthers from the flowers of one parental strain so that it couldn't self-pollinate. Then he collected pollen from the other parental strain and placed it on the stigmas of flowers of the strain whose anthers had been removed. The plants providing and receiving the pollen were the **parental generation**, designated **P**.
- In due course, seeds formed and were planted. The seeds and the resulting new plants constituted the **first filial generation**, or **F₁**. (The word "filial" refers to the relationship between offspring and parents, from the Latin, *filius*, "son.") Mendel and his assistants examined each **F₁** plant to see which traits it bore and then recorded the number of **F₁** plants expressing each trait.

TABLE 12.1
Mendel's Results from Monohybrid Crosses

	PARENTAL GENERATION PHENOTYPES		F ₂ GENERATION PHENOTYPES			TOTAL	RATIO
	DOMINANT	RECESSIVE	DOMINANT	RECESSIVE			
	Spherical seeds × Wrinkled seeds		5,474	1,850	7,324	2.96:1	
	Yellow seeds × Green seeds		6,022	2,001	8,023	3.01:1	
	Purple flowers × White flowers		705	224	929	3.15:1	
	Inflated pods × Constricted pods		882	299	1,181	2.95:1	
	Green pods × Yellow pods		428	152	580	2.82:1	
	Axial flowers × Terminal flowers		651	207	858	3.14:1	
	Tall stems × Dwarf stems (1 m) (0.3 m)		787	277	1,064	2.84:1	

- In some experiments the F₁ plants were allowed to self-pollinate and produce a **second filial generation**, the F₂. Again, each F₂ plant was characterized and counted.

Mendel's first experiments involved monohybrid crosses

The term *hybrid* refers to the offspring of crosses between organisms differing in one or more traits. In Mendel's first experiment, he crossed two true-breeding parental (P) lineages differing in just one trait, producing monohybrids in the F₁ generation. He subsequently planted the F₁ seeds and allowed the resulting plants to self-pollinate to produce the F₂ generation. This technique is referred to as a **monohybrid cross**, even though in this case, the monohybrid plants were not literally crossed, but self-pollinated.

Mendel performed the same experiment for all seven pea traits. His method is illustrated in **Figure 12.3**, using the seed shape trait as an example. He took pollen from pea plants of a true-breeding strain with wrinkled seeds and placed it on the stigmas of flowers of a true-breeding strain with spherical seeds. He also performed the reciprocal cross, in which the parental source of each trait is reversed: he placed pollen from the spherical-seeded strain on the stigmas of flowers of the strain with wrinkled seeds. In all cases, the F₁ seeds were spherical—it was as if the wrinkled seed trait had disappeared completely.

The following spring, Mendel grew 253 F₁ plants from these spherical seeds. Each of the F₁ plants was allowed to self-pollinate to produce F₂ seeds. In all, 7,324 F₂ seeds were produced,

of which 5,474 were spherical, and 1,850 wrinkled (see Figure 12.3 and Table 12.1).

Mendel observed that the wrinkled seed trait was never expressed in the F₁ generation, even though it reappeared in the F₂ generation. This led him to conclude that the spherical seed trait was **dominant** to the wrinkled seed trait, which he called **recessive**. In each of the other six pairs of traits Mendel studied, one trait proved to be dominant over the other trait. The trait that disappears in the F₁ generation of a true-breeding cross is always the recessive trait.

Mendel also observed that the ratio of the two traits in the F₂ generation was always the same—approximately 3:1—for each of the seven pea-plant traits that he studied. That is, three-fourths of the F₂ generation showed the dominant trait and one-fourth showed the recessive trait (see Table 12.1). For example, Mendel's monohybrid cross for seed shape produced a ratio of 5,474:1,850 = 2.96:1. The two reciprocal crosses in the parental generation yielded similar outcomes in the F₂; it did not matter which parent contributed the pollen, just as Kölreuter had shown.

REJECTION OF THE BLENDING THEORY Mendel's monohybrid cross experiments showed that inheritance cannot be the result of a blending phenomenon. According to the blending theory, Mendel's F₁ seeds should have had an appearance that was intermediate between those of the two parents—for example, the F₁ seeds from the cross between strains with wrinkled and spherical seeds should have been slightly wrinkled. Furthermore, the blending theory offered no explanation for the reappearance of the recessive trait in the F₂ seeds after its absence in the F₁ seeds.

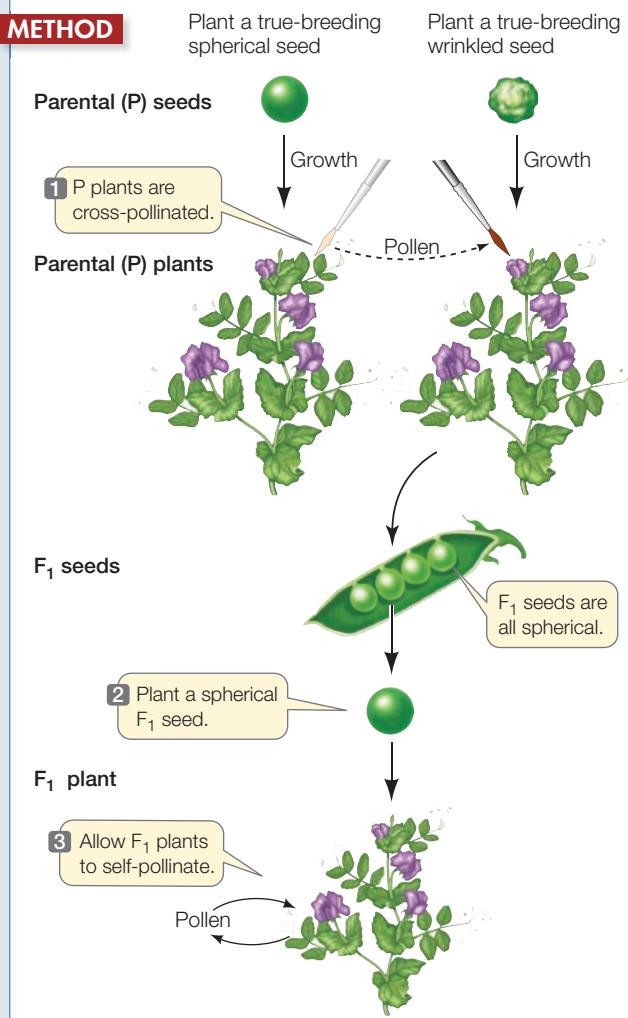
INVESTIGATING LIFE

12.3 Mendel's Monohybrid Experiments

Mendel performed crosses with pea plants and carefully analyzed the outcomes to show that genetic determinants are particulate.

HYPOTHESIS When two strains of peas with contrasting traits are bred, their characteristics are irreversibly blended in succeeding generations.

METHOD



RESULTS

F₂ seeds from F₁ plant

- 4 F₂ seeds: 3/4 are spherical, 1/4 are wrinkled (3:1 ratio).

CONCLUSION

The hypothesis is rejected. There is no irreversible blending of characteristics, and a recessive trait can reappear in succeeding generations.

Go to yourBioPortal.com for original citations, discussions, and relevant links for all INVESTIGATING LIFE figures.

SUPPORT FOR THE PARTICULATE THEORY Given the absence of blending and the reappearance of the recessive seed traits in the F₂ generations of his monohybrid cross experiments, Mendel proposed that the units responsible for the inheritance of specific traits are present as discrete particles that occur in pairs and segregate (separate) from one another during the formation of gametes. According to his **particulate theory**, the units of inheritance retain their integrity in the presence of other units. Mendel concluded that each pea plant has two units (particles) of inheritance for each character, one from each parent. We now use the term **diploid** to refer to the two copies of each heritable unit in an organism. Mendel proposed that during the production of gametes, only one of these paired units is given to a gamete. We now use the term **haploid** to refer to the single set of heritable units. Mendel concluded that while each gamete contains one unit, the resulting zygote contains two, because it is produced by the fusion of two gametes. This conclusion is the core of Mendel's model of inheritance. Mendel's unit of inheritance is now called a **gene**. The totality of all the genes of an organism is that organism's **genome**.

Mendel reasoned that in his experiments, the two true-breeding parent plants had different forms of the gene affecting a particular character, such as seed shape (although he did not use the term "gene"). The true-breeding spherical-seeded parent had two genes of the same form, which we will call S, and the parent with wrinkled seeds had two copies of an alternative form of the gene, which we will call s. The SS parent would produce gametes having a single S gene, and the ss parent would produce gametes having a single s gene. The cross producing the F₁ generation would donate an S from one parent and an s from the other to each seed; the F₁ offspring would thus be Ss. We say that S is dominant over s because the trait specified by s is not evident—*is not expressed*—when both forms of the gene are present.

Alleles are different forms of a gene

The different forms of a gene (S and s in this case) are called **alleles**. Individuals that are true-breeding for a trait contain two copies of the same allele. For example, all the individuals in a population of true-breeding peas with wrinkled seeds must have the allele pair ss; if the dominant S allele were present, some of the plants would produce spherical seeds.

We say that the individuals that produce wrinkled seeds are **homozygous** for the allele s, meaning that they have two copies of the same allele (ss). Some peas with spherical seeds—the ones with the genotype SS—are also homozygous. However, not all plants with spherical seeds have the SS genotype. Some spherical-seeded plants, like Mendel's F₁, are **heterozygous**: they have two different alleles of the gene in question (in this case, Ss). An individual that is homozygous for a character is sometimes called a **homozygote**; an individual that is heterozygous for a character is termed a **heterozygote**.

As a somewhat more complex example of inheritance, let's consider three gene pairs. An individual with the three genes and alleles *AABbcc* is homozygous for the *A* and *C* genes, because it has two *A* alleles and two *c* alleles, but heterozygous for the *B* gene, because it contains the *B* and *b* alleles.

The physical appearance of an organism is its **phenotype**. Mendel correctly supposed the phenotype to be the result of the **genotype**, or genetic constitution, of the organism showing the phenotype. Spherical seeds and wrinkled seeds are two phenotypes, which are the result of three genotypes: the wrinkled seed phenotype is produced by the genotype *ss*, whereas the spherical seed phenotype is produced by either of the genotypes *SS* or *Ss*.

Mendel's first law says that the two copies of a gene segregate

How does Mendel's model of inheritance explain the ratios of traits seen in the F_1 and F_2 generations? Consider first the F_1 , in which all progeny have the spherical seed phenotype. According to Mendel's model, *when any individual produces gametes, the two copies of a gene separate, so that each gamete receives only one copy*. This is Mendel's first law, the **law of segregation**. Thus, every individual in the offspring from a cross between the P generation parents inherits one gene copy from each parent, and has the genotype *Ss* (Figure 12.4).

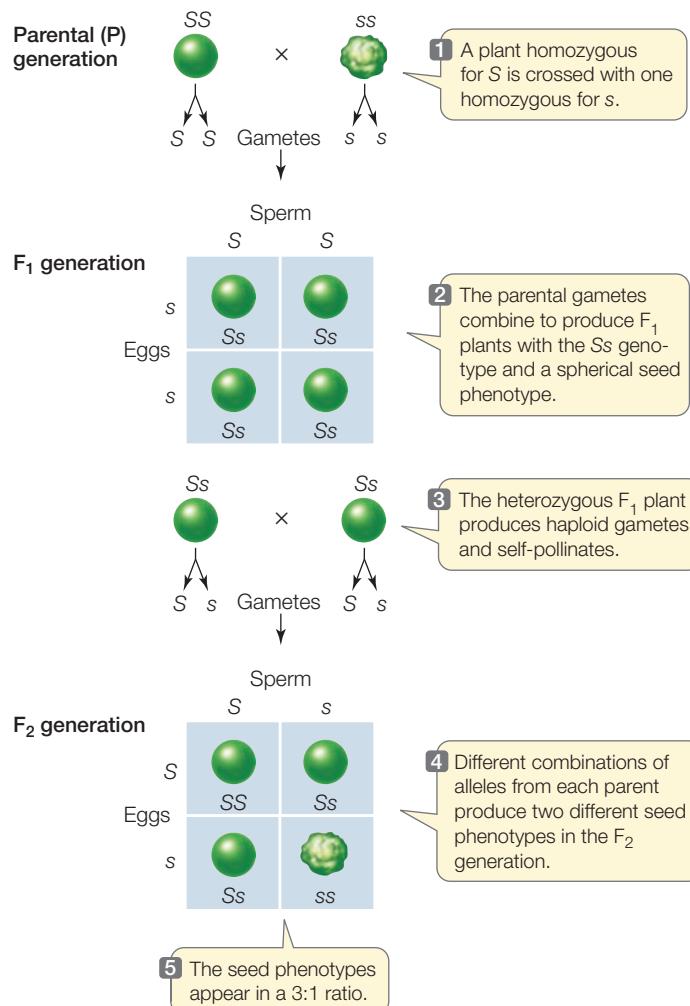
Now let's consider the composition of the F_2 generation. Half of the gametes produced by the F_1 generation have the *S* allele and the other half the *s* allele. Since both *SS* and *Ss* plants produce spherical seeds while *ss* plants produce wrinkled seeds, in the F_2 generation there are three ways to get a spherical-seeded plant (*SS*, *Ss*, or *sS*), but only one way to get a plant with wrinkled seeds (*ss*). Therefore, we predict a 3:1 ratio, remarkably close to the values Mendel found experimentally for all seven of the traits he compared (see Table 12.1).

The allele combinations that will result from a cross can be predicted using a **Punnett square**, a method devised in 1905 by the British geneticist Reginald Crundall Punnett. This device ensures that we consider all possible combinations of gametes when calculating expected genotype frequencies. A Punnett square looks like this:

		Male gametes	
		<i>S</i>	<i>s</i>
Female gametes	<i>S</i>	<i>SS</i>	<i>Ss</i>
	<i>s</i>	<i>Ss</i>	<i>ss</i>

It is a simple grid with all possible male gamete (haploid sperm) genotypes shown along the top and all possible female gamete (haploid egg) genotypes along the left side. The grid is completed by filling in each square with the diploid genotype that can be generated from each combination of gametes (see Figure 12.4). In this example, to fill in the top right square, we put in the *S* from the egg (female gamete) and the *s* from the pollen (male gamete), yielding *Ss*.

Mendel did not live to see his theory placed on a sound physical footing with the discoveries of chromosomes and DNA. Genes are now known to be regions of the DNA molecules in chromosomes. More specifically, a gene is a sequence of DNA that resides at a particular site on a chromosome, called a **locus** (plural **loci**). Genes are expressed in the phenotype mostly as proteins with particular functions, such as enzymes.

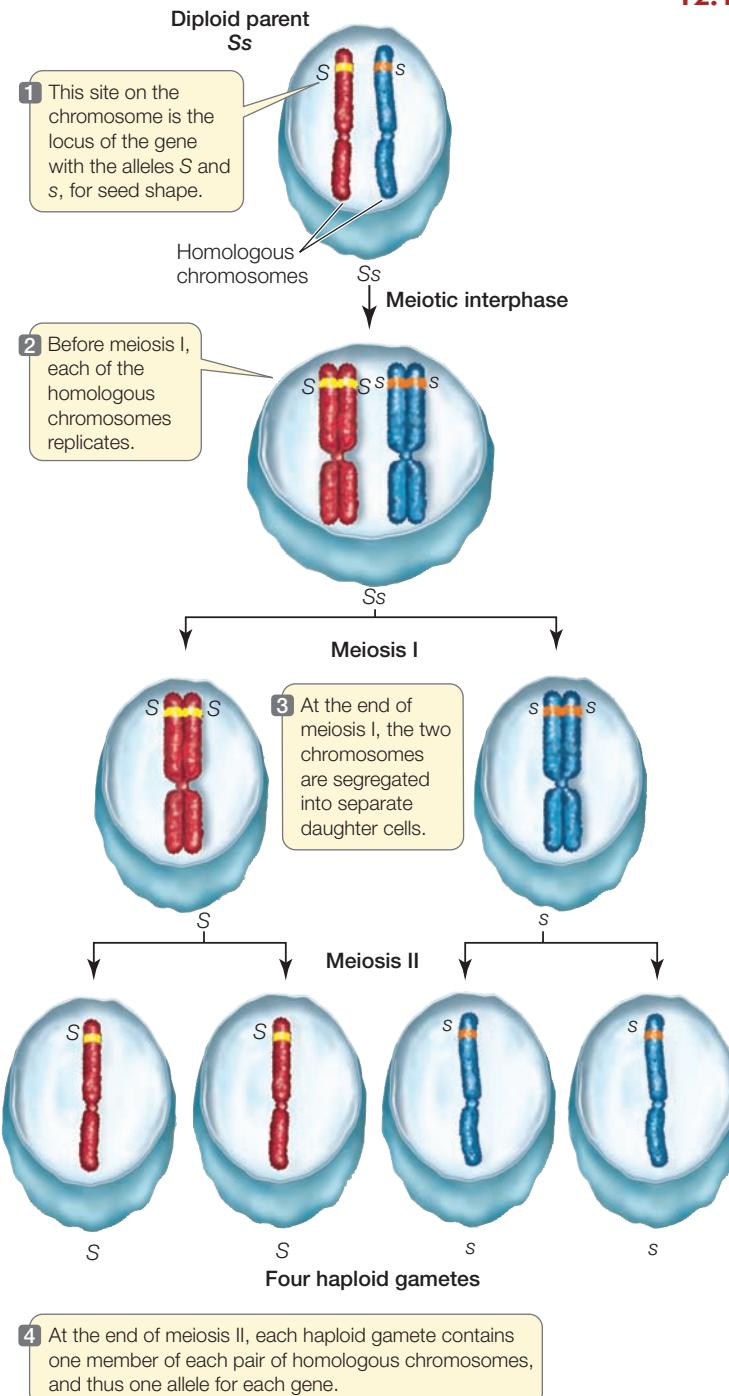


12.4 Mendel's Explanation of Inheritance Mendel concluded that inheritance depends on discrete factors from each parent that do not blend in the offspring.

So, in many cases, a dominant gene can be thought of as a region of DNA that is expressed as a functional protein, while a recessive gene typically expresses a nonfunctional protein, or a protein whose function is overshadowed by the dominant form. Mendel arrived at his law of segregation with no knowledge of chromosomes or meiosis, but today we can picture the different alleles of a gene segregating as the chromosomes separate during meiosis I (Figure 12.5).

Mendel verified his hypothesis by performing a test cross

Mendel set out to test his hypothesis that there were two possible allele combinations (*SS* and *Ss*) in the spherical-seeded F_1 generation. He did so by performing a **test cross**, which is a way of finding out whether an individual showing a dominant trait is homozygous or heterozygous. In a test cross, the individual in question is crossed with an individual that is known to be homozygous for the recessive trait—an easy individual to identify, because all individuals with the recessive phenotype are homozygous for the recessive trait.



12.5 Meiosis Accounts for the Segregation of Alleles Although Mendel had no knowledge of chromosomes or meiosis, we now know that a pair of alleles resides on homologous chromosomes, and that those alleles segregate during meiosis.

For the seed shape gene that we have been considering, the recessive homozygote used for the test cross is ss . The individual being tested may be described initially as $S-$ because we do not yet know the identity of the second allele. We can predict two possible results:

- If the individual being tested is homozygous dominant (SS), all offspring of the test cross will be Ss and show the dominant trait (spherical seeds) (Figure 12.6, left).

INVESTIGATING LIFE

12.6 Homozygous or Heterozygous?

An individual with a dominant phenotype may have either a homozygous or a heterozygous genotype. The test cross determines which.

HYPOTHESIS The progeny of a test cross can reveal whether an organism is homozygous or heterozygous.

METHOD

1a Test spherical peas of undetermined genotype... $S-$ \times ss

1b ...by crossing them with wrinkled peas with a known genotype (homozygous recessive).

2a If the plant being tested is homozygous... $SS \times ss$ Gametes: S , S ; s , s

2b If the plant being tested is heterozygous... $Ss \times ss$ Gametes: S , s ; s , s

RESULTS

		Sperm	
		S	s
Eggs	S	Ss	Ss
	s	Ss	Ss

3a ...then all progeny will show the dominant phenotype (spherical).

		Sperm	
		S	s
Eggs	S	Ss	Ss
	s	ss	ss

3b ...then half the seeds from the cross will be wrinkled, and half will be spherical.

CONCLUSION

The plant being tested is homozygous.

CONCLUSION

The plant being tested is heterozygous.

FURTHER INVESTIGATIONS: What would be the result if the "tester" plant was homozygous for spherical instead of wrinkled seeds?

Go to yourBioPortal.com for original citations, discussions, and relevant links for all INVESTIGATING LIFE figures.

- If the individual being tested is heterozygous (Ss), then approximately half of the offspring of the test cross will be heterozygous and show the dominant trait (Ss), but the other half will be homozygous for, and will show, the recessive trait (ss) (Figure 12.6, right).

Mendel obtained results consistent with both of these predictions; thus Mendel's hypothesis accurately predicted the results of his test crosses.

With his first hypothesis confirmed, Mendel went on to ask another question: How do different pairs of genes behave in crosses when considered together?

Mendel's second law says that copies of different genes assort independently

Consider an organism that is heterozygous for two genes ($SsYy$), in which the S and Y alleles came from its mother, and the s and y alleles came from its father. When this organism produces gametes, do the alleles of maternal origin (S and Y) go together in one gamete and those of paternal origin (s and y) in another gamete? Or can a single gamete receive one maternal and one paternal allele, S and y (or s and Y)?

To answer these questions, Mendel performed another series of experiments. He began with peas that differed in two seed characters: seed shape and seed color. One true-breeding parental strain produced only spherical, yellow seeds ($SSYY$), and the other produced only wrinkled, green ones ($ssyy$). A cross between these two strains produced an F_1 generation in which all the plants were $SsYy$. Because the S and Y alleles are dominant, the F_1 seeds were all spherical and yellow.

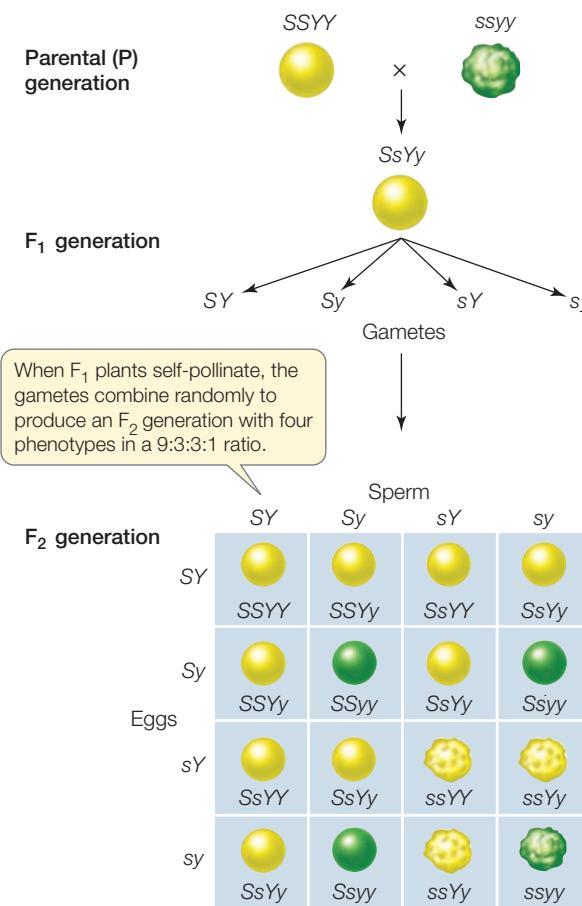
Mendel continued this experiment into the F_2 generation by performing a **dihybrid cross** (a cross between individuals that are identical double heterozygotes) with the F_1 plants (although again, in this case, this was done by allowing the F_1 plants to self-pollinate). There are two possible ways in which such doubly heterozygous plants might produce gametes, as Mendel saw it (remember that he had never heard of chromosomes or meiosis):

1. *The alleles could maintain the associations they had in the parental generation (that is, they could be linked).*

In this case, the F_1 plants should produce two types of gametes (SY and sy), and the F_2 progeny resulting from self-pollination of the F_1 plants should consist of three times as many plants bearing spherical, yellow seeds as plants with wrinkled, green seeds. If such results were obtained, there might be no reason to suppose that two different genes regulated seed shape and seed color, because spherical seeds would always be yellow and wrinkled ones always green.

2. *The segregation of S from s could be independent of the segregation of Y from y (that is, the two genes could be unlinked).*

In this case, four kinds of gametes should be produced by the F_1 in equal numbers: SY , Sy , sY , and sy . When these gametes combine at random, they should produce an F_2 having nine different genotypes. The F_2 progeny could have any of three possible genotypes for shape (SS , Ss , or ss) and any of three possi-

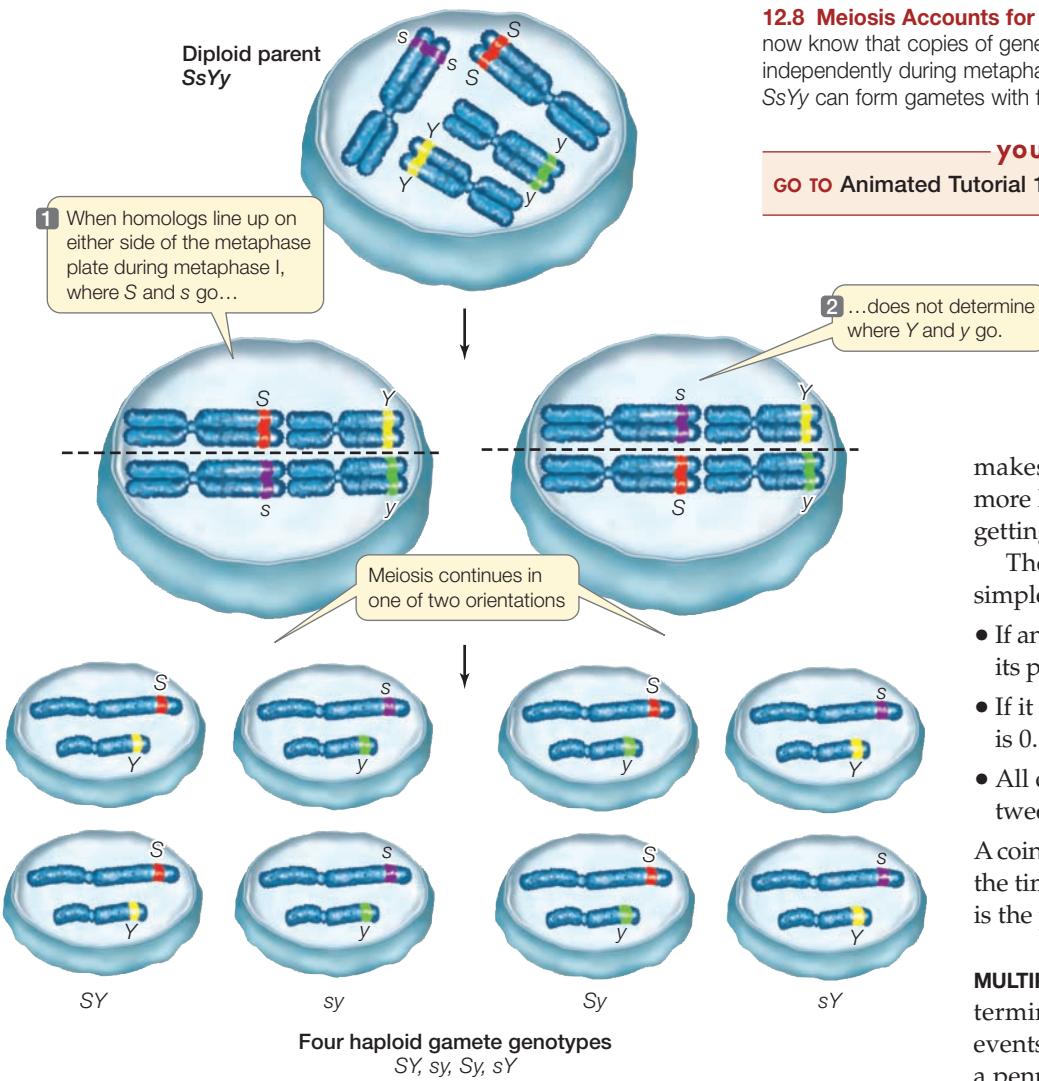


12.7 Independent Assortment The 16 possible combinations of gametes in this dihybrid cross result in nine different genotypes. Because S and Y are dominant over s and y , respectively, the nine genotypes result in four phenotypes in a ratio of 9:3:3:1. These results show that the two genes segregate independently.

ble genotypes for color (YY , Yy , or yy). The combined nine genotypes should produce four phenotypes (spherical yellow, spherical green, wrinkled yellow, wrinkled green). Putting these possibilities into a Punnett square, we can predict that these four phenotypes will occur in a ratio of 9:3:3:1 (Figure 12.7).

Mendel's dihybrid crosses supported the second prediction: four different phenotypes appeared in the F_2 generation in a ratio of about 9:3:3:1. The parental traits appeared in new combinations (spherical green and wrinkled yellow) in some progeny. Such new combinations are called **recombinant** phenotypes.

These results led Mendel to the formulation of what is now known as Mendel's second law: *alleles of different genes assort independently of one another during gamete formation*. That is, the segregation of gene A alleles is independent of the segregation of gene B alleles. We now know that this **law of independent assortment** is not as universal as the law of segregation, because it applies to genes located on separate chromosomes, but not to those located near one another on the same chromosome, as we will see in Section 12.4. However, it is correct to say that chromosomes segregate independently during the formation of gametes, and so do any two genes on separate homologous chromosome pairs (Figure 12.8).



One of Mendel's major contributions to the science of genetics was his use of the rules of statistics and probability to analyze his masses of data from hundreds of crosses resulting in thousands of progeny plants. His mathematical analyses revealed clear patterns in the data that allowed him to formulate his hypotheses. Ever since his work became widely recognized, geneticists have used simple mathematics in the same ways that Mendel did.

Punnett squares or probability calculations: A choice of methods

Punnett squares provide one way of solving problems in genetics, and probability calculations provide another. Many people find it easier to use the principles of probability, some of which are intuitive and familiar. For example, when we flip a coin, the law of probability states that it has an equal probability of landing "heads" or "tails." For any given toss of a fair coin, the probability of heads is independent of what happened in all the previous tosses. A run of ten straight heads implies nothing about the next toss. No "law of averages" increases the likelihood that the next toss will come up tails, and no "momentum"

12.8 Meiosis Accounts for Independent Assortment of Alleles We now know that copies of genes on different chromosomes are segregated independently during metaphase I of meiosis. Thus a parent of genotype $SsYy$ can form gametes with four different genotypes.

yourBioPortal.com
[GO TO](#) Animated Tutorial 12.1 • Independent Assortment of Alleles

makes an eleventh occurrence of heads any more likely. On the eleventh toss, the odds of getting heads are still 50-50.

The basic conventions of probability are simple:

- If an event is absolutely certain to happen, its probability is 1.
- If it cannot possibly happen, its probability is 0.
- All other events have a probability between 0 and 1.

A coin toss results in heads approximately half the time, so the probability of heads is $\frac{1}{2}$ —as is the probability of tails.

MULTIPLYING PROBABILITIES How can we determine the probability of two independent events happening together? If two coins (say a penny and a dime) are tossed, each acts independently of the other. What is the probability of both coins coming up heads? In half of the tosses, the penny comes up heads; of that fraction, the dime also comes up heads half of the time. Therefore, the joint probability of both coins coming up heads is half of one-half, or $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$. So, to find the joint probability of independent events, we multiply the probabilities of the individual events (Figure 12.9). How does this method apply to genetics?

To see how joint probability is calculated in genetics problems, let's consider the monohybrid cross. The probabilities of two events are involved: gamete formation and random fertilization.

Calculating the probabilities involved in gamete formation is straightforward. A homozygote can produce only one type of gamete, so, for example, the probability of an SS individual producing gametes with the genotype S is 1. The heterozygote Ss produces S gametes with a probability of $\frac{1}{2}$ and s gametes with a probability of $\frac{1}{2}$.

Now let's see how the rules of probability might predict the ratio of the F_2 progeny of the cross shown in Figure 12.4. These plants are obtained by the self-pollination of F_1 plants of genotype Ss . The probability that an F_2 plant will have the genotype SS must be $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$, because there is a 50-50 chance that the

12.9 Using Probability Calculations in Genetics Like the results of a coin toss, the probability of any given combination of alleles appearing in the offspring of a cross can be obtained by multiplying the probabilities of each event. Since a heterozygote can be formed in two ways, these two probabilities are added together.

sperm will have the genotype *S*, and an independent chance of 50–50 that the egg will have the genotype *S*. Similarly, the probability of *ss* offspring is also $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$.

ADDING PROBABILITIES How are probabilities calculated when an event can happen in different ways? The probability of an *F*₂ plant getting an *S* allele from the sperm and an *s* allele from the egg is $\frac{1}{4}$. In addition, there is a probability of $\frac{1}{4}$ that the *F*₂ plant will get an *s* from the sperm and an *S* from the egg, resulting in the same genotype of *Ss*. The probability of an event that can occur in two or more different ways is the sum of the individual probabilities of those ways. Thus the probability that an *F*₂ plant will be a heterozygote is equal to the sum of the probabilities of the two ways of forming a heterozygote: $\frac{1}{4} + \frac{1}{4} = \frac{1}{2}$ (see Figure 12.9). The three genotypes are therefore expected to occur in the ratio $\frac{1}{4} SS : \frac{1}{2} Ss : \frac{1}{4} ss$, resulting in the 1:2:1 ratio of genotypes and the 3:1 ratio of phenotypes seen in Figure 12.4.

PROBABILITY AND THE DIHYBRID CROSS If *F*₁ plants heterozygous for two independent characters self-pollinate, the resulting *F*₂ plants express four different phenotypes. The proportions of these phenotypes are easily determined by probability calculations. Let's see how this works for the experiment shown in Figure 12.7.

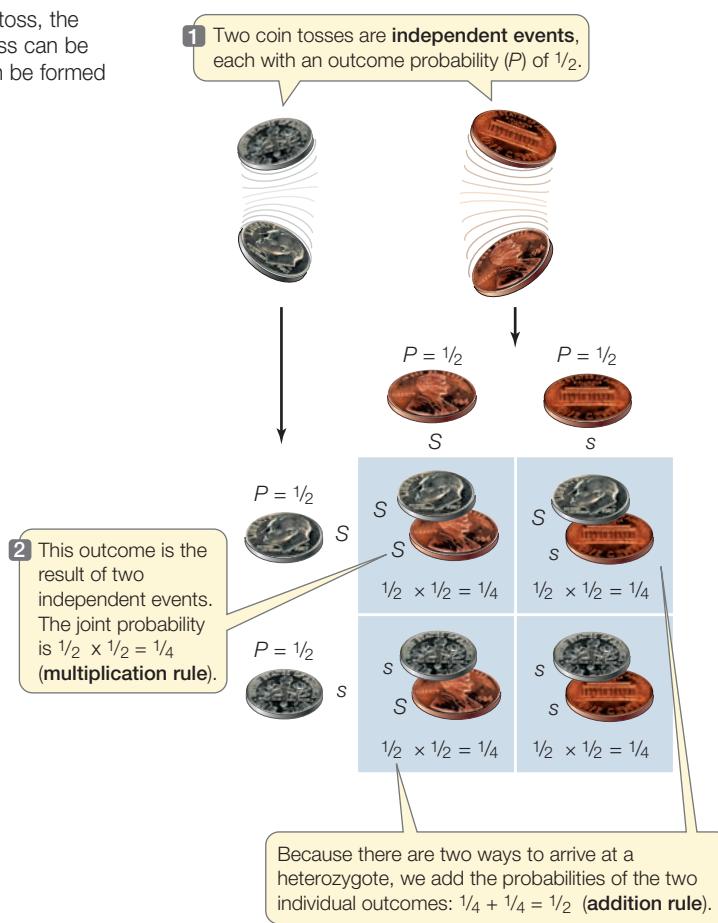
Using the principles described above, we can calculate that the probability that an *F*₂ seed will be spherical is $\frac{3}{4}$. This is found by adding the probability of an *Ss* heterozygote ($\frac{1}{2}$) and the probability of an *SS* homozygote ($\frac{1}{4}$) = a total of $\frac{3}{4}$. By the same reasoning, the probability that a seed will be yellow is also $\frac{3}{4}$. The two characters are determined by separate genes and are independent of each other, so the joint probability that a seed will be both spherical and yellow is $\frac{3}{4} \times \frac{3}{4} = \frac{9}{16}$. What is the probability of *F*₂ seeds being both wrinkled and yellow? The probability of being yellow is again $\frac{3}{4}$; the probability of being wrinkled is $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$. The joint probability that a seed will be *both* wrinkled and yellow is $\frac{1}{4} \times \frac{3}{4} = \frac{3}{16}$.

The same probability applies, for similar reasons, to spherical, green *F*₂ seeds. Finally, the probability that *F*₂ seeds will be both wrinkled and green is $\frac{1}{4} \times \frac{1}{4} = \frac{1}{16}$. Looking at all four phenotypes, we see that they are expected to occur in the ratio of 9:3:3:1.

Probability calculations and Punnett squares give the same results. Learn to do genetics problems both ways, and then decide which method you prefer.

Mendel's laws can be observed in human pedigrees

How are Mendel's laws of inheritance applied to humans? Mendel worked out his laws by performing many planned



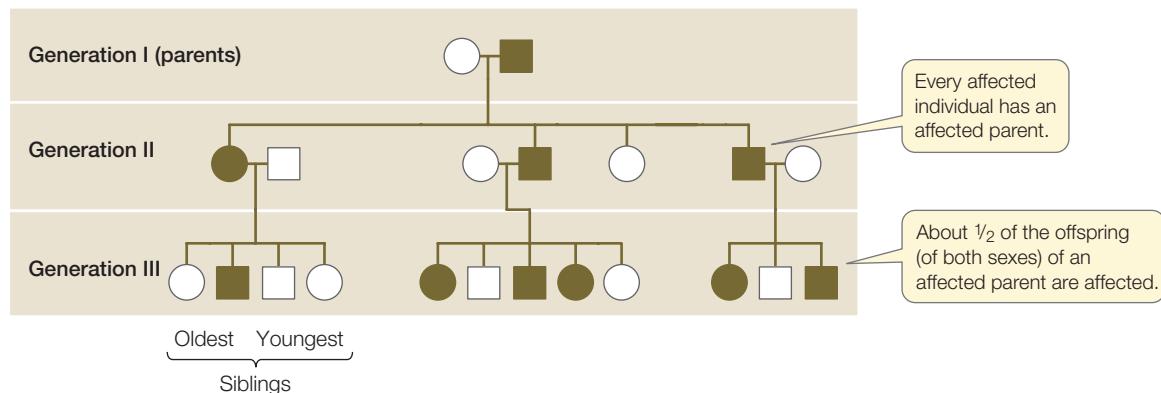
crosses and counting many offspring. Neither of these approaches is possible with humans, so human geneticists rely on **pedigrees**: family trees that show the occurrence of phenotypes (and alleles) in several generations of related individuals.

Because humans have such small numbers of offspring, human pedigrees do not show the clear proportions of offspring phenotypes that Mendel saw in his pea plants. For example, when a man and a woman who are both heterozygous for a recessive allele (say, *Aa*) have children together, each child has a 25 percent probability of being a recessive homozygote (*aa*). Thus if this couple were to have dozens of children, one-fourth of them would be recessive homozygotes. But the offspring of a single couple are likely to be too few to show the exact one-fourth proportion. In a family with only two children, for example, both could easily be *aa* (or *Aa*, or *AA*).

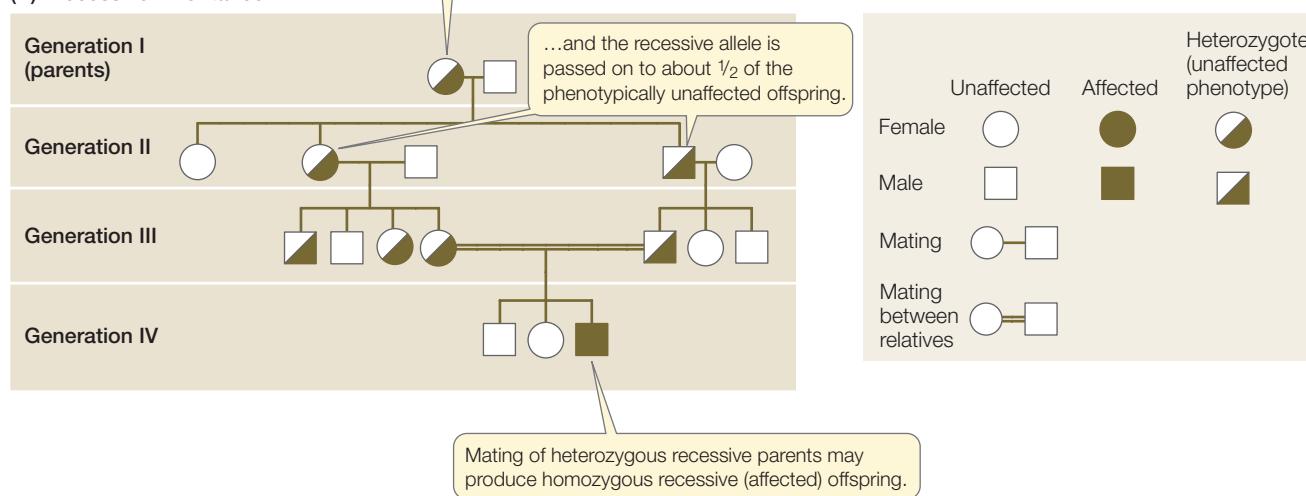
What if we want to know whether a recessive allele is carried by both the mother and the father? Human geneticists assume that any allele that causes an abnormal phenotype (such as a genetic disease) is rare in the human population. This means that if some members of a given family have a rare allele, it is highly unlikely that an outsider marrying into that family will have that same rare allele.

Human geneticists may wish to know whether a particular rare allele that causes an abnormal phenotype is dominant or recessive. **Figure 12.10A** is a pedigree showing the pattern of

(A) Dominant inheritance



(B) Recessive inheritance



12.10 Pedigree Analysis and Inheritance (A) This pedigree represents a family affected by Huntington's disease, which results from a rare dominant allele. Everyone who inherits this allele is affected. (B) The family in this pedigree carries the allele for albinism, a recessive trait. Because the trait is recessive, heterozygotes do not have the albino phenotype, but they can pass the allele on to their offspring. Affected persons must inherit the allele from two heterozygous parents, or (rarely) from one homozygous recessive and one heterozygous parent, or (very rarely) two homozygous recessive parents. In this family, in generation III the heterozygous parents are cousins; however, the same result could occur if the parents were unrelated but heterozygous.

inheritance of a rare dominant allele. The following are the key features to look for in such a pedigree:

- Every affected person has an affected parent.
- About half of the offspring of an affected parent are also affected.
- The phenotype occurs equally in both sexes.

Compare this pattern with the one shown in **Figure 12.10B**, which is typical for the inheritance of a rare recessive allele:

- Affected people usually have two parents who are not affected.

- In affected families, about one-fourth of the children of unaffected parents are affected.

- The phenotype occurs equally in both sexes.

In pedigrees showing inheritance of a recessive phenotype, it is not uncommon to find a marriage of two relatives. This observation is a result of the rarity of recessive alleles that give rise to abnormal phenotypes. For two phenotypically normal parents to have an affected child (aa), the parents must both be heterozygous (Aa). If a particular recessive allele is rare in the general population, the chance of two people marrying who are both carrying that allele is quite low. On the other hand, if that allele is present in a family, two cousins might share it (see Figure 12.10B). For this reason, studies on populations that are isolated either culturally (by religion, as with the Amish in the United States) or geographically (as on islands) have been extremely valuable to human geneticists. People in these groups are more likely to marry relatives who may carry the same rare recessive alleles.

Because the major use of pedigree analysis is in the clinical evaluation and counseling of patients with inherited abnormalities in their families, a single pair of alleles is usually followed. However, pedigree analysis can also show independent assortment if two different allele pairs are considered.

12.1 RECAP

Mendel showed that genetic determinants are particulate and do not “blend” or disappear when the genes from two gametes combine. Mendel’s first law of inheritance states that the two copies of a gene segregate during gamete formation. His second law states that genes assort independently during gamete formation. The frequencies with which different allele combinations will be expressed in offspring can be calculated with a Punnett square or using probability theory.

- What results seen in the F_1 and F_2 generations of Mendel’s monohybrid cross experiments refuted the blending theory of inheritance? **See p. 240, Figures 12.3 and 12.4, and Table 12.1**
- How do events in meiosis explain Mendel’s monohybrid cross results? **See pp. 242–244 and Figure 12.5**
- How do events in meiosis explain the independent assortment of alleles in Mendel’s dihybrid cross experiments? **See p. 244 and Figures 12.7 and 12.8**
- Draw human pedigrees for dominant and recessive inheritance. **See pp. 246–247 and Figure 12.10**

The laws of inheritance as articulated by Mendel remain valid today; his discoveries laid the groundwork for all future studies of genetics. Inevitably, however, we have learned that things are more complicated. Let’s take a look at some of these complications, beginning with the interactions between alleles at different loci.

12.2 How Do Alleles Interact?

Existing alleles are subject to change, and thus may give rise to new alleles, so there can be many alleles for a single character. In addition, alleles do not always show simple dominant-recessive relationships. Furthermore, a single allele may have multiple phenotypic effects.

Possible genotypes	CC, Cc^{ch}, Cc^h, Cc	$c^{ch}c^{ch}$	$c^{ch}c^h, c^{ch}c$	$c^h c^h, c^h c$	cc
Phenotype	Dark gray	Chinchilla	Light gray	Point restricted	Albino



New alleles arise by mutation

Genes are subject to **mutations**, which are rare, stable, and inherited changes in the genetic material. In other words, an allele can mutate to become a different allele. For example, you can envision that at one time all pea plants were tall and had the height allele T . A mutation occurred in that allele that resulted in a new allele, t (short). If this mutation was in a cell that underwent meiosis to form gametes, some of the resulting gametes would carry the t allele, and some offspring of this pea plant would carry the t allele. Mutation will be discussed in detail in Chapter 15. By creating variety, mutations are the raw material for evolution.

Geneticists usually define one particular allele of a gene as the **wild type**; this allele is the one that is present in most individuals in nature (“the wild”) and gives rise to an expected trait or phenotype. Other alleles of that gene, often called mutant alleles, may produce a different phenotype. The wild-type and mutant alleles reside at the same locus and are inherited according to the rules set forth by Mendel. A genetic locus with a wild-type allele that is present less than 99 percent of the time (the rest of the alleles being mutant) is said to be **polymorphic** (Greek *poly*, “many”; *morph*, “form”).

Many genes have multiple alleles

Because of random mutations, more than two alleles of a given gene may exist in a group of individuals. (Any one individual has only two alleles—one from its mother and one from its father. But different individuals may carry several different alleles.) In fact, there are many examples of such multiple alleles, and they often show a hierarchy of dominance.

Coat color in rabbits, for example, is determined by one gene with four alleles:

- C determines dark gray
- c determines albino
- c^{ch} determines chinchilla
- c^h determines light gray

12.11 Inheritance of Coat Color in Rabbits There are four alleles of the gene for coat color in these Netherlands dwarf rabbits. Different combinations of two alleles give different coat colors. The dominance hierarchy is $C > c^{ch} > c^h > c$.

Any rabbit with the C allele (paired with any of the four) is dark gray, and a rabbit with *cc* is albino. The intermediate colors result from the different allele combinations shown in **Figure 12.11**.

Multiple alleles increase the number of possible phenotypes. Each of Mendel's monohybrid crosses involved just one pair of alleles (for example, *S* and *s*) and two possible phenotypes (resulting from *SS* or *Ss* and *ss*). The four alleles of the rabbit coat color gene produce five different phenotypes.

Dominance is not always complete

In the pairs of alleles studied by Mendel, dominance is complete when an individual is heterozygous. That is, an *Ss* individual always expresses the *S* phenotype. However, many genes have alleles that are not dominant or recessive to one another. Instead, the heterozygotes show an intermediate phenotype—at first glance, like that predicted by the old blending theory of inheritance. For example, if a true-breeding red snapdragon is crossed with a true-breeding white one, all the *F*₁ flowers are pink. However, further crosses indicate that this apparent blending phenomenon can still be explained in terms of Mendelian genetics (**Figure 12.12**). The red and white alleles have not disappeared, as those colors reappear when the *F*₁ plants are interbred.

We can understand these results in terms of the Mendelian laws of inheritance. When heterozygotes show a phenotype that is intermediate between those of the two homozygotes, the gene is said to be governed by **incomplete dominance**. In other words, neither of the two alleles is dominant. Incomplete dominance is common in nature, and at the biochemical level, most examples of incomplete dominance are actually codominance (see below). In fact, Mendel's study of seven pea-plant traits is unusual in that all seven traits happened to be characterized by complete dominance.

In codominance, both alleles at a locus are expressed

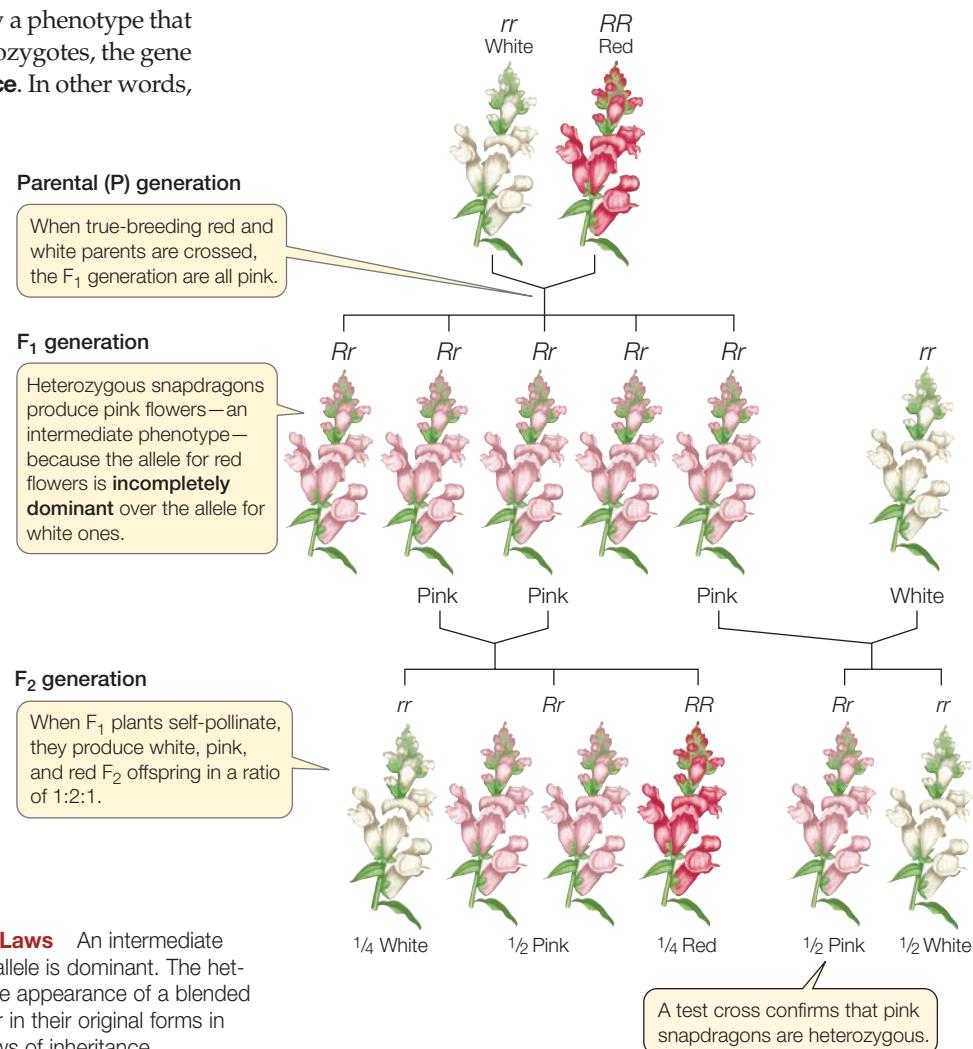
Sometimes the two alleles at a locus produce two different phenotypes that *both* appear in heterozygotes, a phenomenon called **codominance**. Note that this is different from incomplete dominance, where the phenotype of a heterozygote is a blend of the phenotypes of the parents. A good example of codominance is seen in the ABO blood group system in humans.

There are numerous glycoproteins on the surfaces of red blood cells and they are all encoded by genes. One genetic locus is

called the ABO locus, with three alleles, *I^A*, *I^B* and *I^O*, that encode variants of a surface glycoprotein designated A, B, and O (the “ABO system”). Since people inherit one allele from each parent, they may have any combination of these alleles: *I^AI^B*, *I^AI^O*, *I^BI^O*, and so on. In terms of gene expression, it is important to note that in a codominant system, all alleles are expressed in a heterozygote. So people with *I^AI^B* express both *I^A* and *I^B* alleles on their red blood cell surfaces.

Early attempts at blood transfusion frequently killed the patient. Around 1900, the Austrian scientist Karl Landsteiner mixed blood cells and serum (blood from which cells have been removed) from different individuals. He found that only certain combinations of blood and serum are compatible. In other combinations, the red blood cells from one individual form clumps in the presence of serum from the other individual. This discovery led to our ability to administer compatible blood transfusions that do not kill the recipient.

Incompatible transfusions result in the formation of clumps because of genetic systems like the ABO locus. People make specific proteins in the serum, called antibodies, that react with foreign, or “nonself,” molecules called antigens. The A and B glycoproteins can act as antigens if present on the surfaces of red



12.12 Incomplete Dominance Follows Mendel's Laws An intermediate phenotype can occur in heterozygotes when neither allele is dominant. The heterozygous phenotype (here, pink flowers) may give the appearance of a blended trait, but the traits of the parental generation reappear in their original forms in succeeding generations, as predicted by Mendel's laws of inheritance.

blood cells in donated blood. If the person receiving the blood does not carry the I^A or I^B alleles, their antibodies will react with the nonself glycoproteins and the red blood cells will form clumps. The O glycoprotein does not act as an antigen. You can see these relationships in **Figure 12.13**. We will learn much more about the functions of antibodies and antigens in Chapter 42.

Interestingly, a recent development may make it possible to circumvent the ABO system of blood incompatibility. Enzymes have been isolated from bacteria that can convert the A and B glycoproteins into O glycoprotein. So blood from any genotype in the ABO system could be treated with these enzymes to make O-type blood, which is not antigenic. Since I^O is not a common allele in most human populations, this technology may be important in overcoming shortages of genetically suitable blood for transfusions.

Some alleles have multiple phenotypic effects

Mendel's principles were further extended when it was discovered that a single allele can influence more than one phenotype. When a single allele has more than one distinguishable phenotypic effect, we say that the allele is **pleiotropic**. A familiar example of pleiotropy involves the allele responsible for the coloration pattern (light body, darker extremities) of Siamese cats. The same allele is also responsible for the characteristic crossed eyes of Siamese cats. Although these effects appear to be unrelated, both are caused by the protein encoded by this allele.

12.13 ABO Blood Reactions Are Important in Transfusions This table shows the results of mixing red blood cells of types A, B, AB, and O with serum containing anti-A or anti-B antibodies. As you look down the columns, note that each of the types, when mixed separately with anti-A and with anti-B, gives a unique pair of results; this is the basic method by which blood is typed. People with type O blood are good blood donors because O cells do not react with either anti-A or anti-B antibodies. People with type AB blood are good recipients, since they make neither type of antibody. When blood transfusions are incompatible, the reaction (clumping of red blood cells) can have severely adverse consequences for the recipient.

Blood type of cells	Genotype	Antibodies made by body	Reaction to added antibodies	
			Anti-A	Anti-B
A	$I^A I^A$ or $I^A I^O$	Anti-B		
B	$I^B I^B$ or $I^B I^O$	Anti-A		
AB	$I^A I^B$	Neither anti-A nor anti-B		
O	$i^o i^o$	Both anti-A and anti-B		

12.2 RECAP

Genes are subject to random mutations that give rise to new alleles; thus many genes have more than two alleles within a population. Dominance is not necessarily an all-or-nothing phenomenon.

- How does the experiment in Figure 12.12 demonstrate incomplete dominance? [See p. 249](#)
- Explain how blood type AB results from codominance. [See pp. 249–250 and Figure 12.13](#)

Thus far we have treated the phenotype of an organism, with respect to a given character, as a simple result of the alleles of a single gene. In many cases, however, several genes interact to determine a phenotype. To complicate things further, the physical environment may interact with the genetic constitution of an individual in determining the phenotype.

12.3 How Do Genes Interact?

We have just seen how two alleles of the same gene can interact to produce a phenotype. If you consider most complex phenotypes, such as human height, you will realize that they are influenced by the products of many genes. We now turn to the genetics of such gene interactions.

Epistasis occurs when the phenotypic expression of one gene is affected by another gene. For example, two genes (B and E) determine coat color in Labrador retrievers:

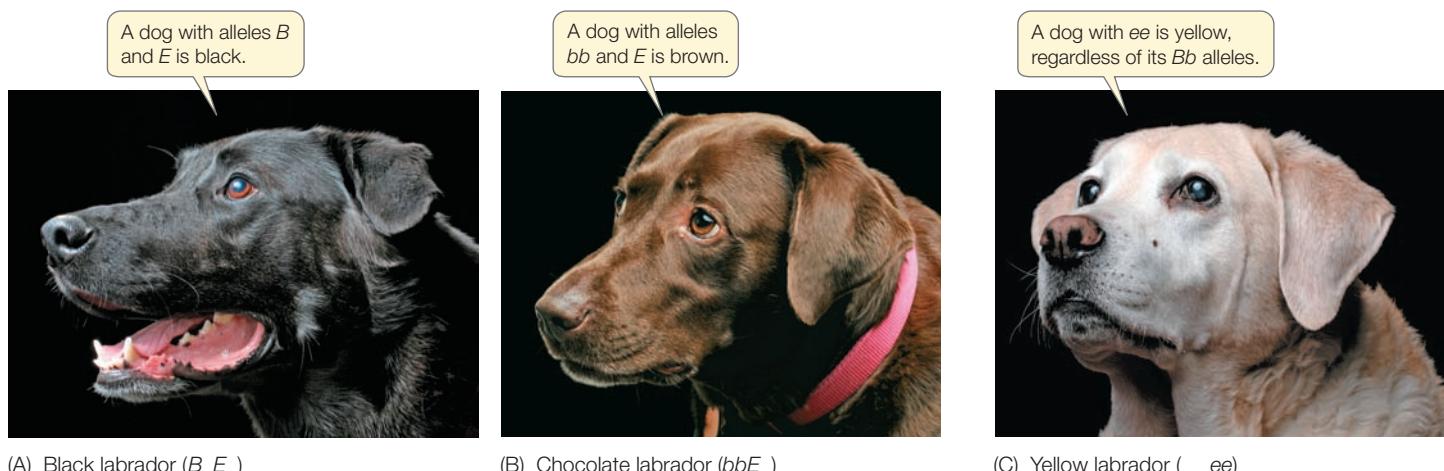
- Allele B (black pigment) is dominant to b (brown)
- Allele E (pigment deposition in hair) is dominant to e (no deposition, so hair is yellow)

So an EE or Ee dog with BB or Bb is black; one with bb is brown; and one with ee is yellow regardless of the Bb alleles present. Clearly, gene E determines the expression of Bb (**Figure 12.14**).

Hybrid vigor results from new gene combinations and interactions

In 1876, Charles Darwin reported that when he crossed two different true-breeding, homozygous genetic strains of corn, the offspring were 25 percent taller than either of the parent strains. Darwin's observation was largely ignored for the next 30 years. In 1908, George Shull "rediscovered" this idea, reporting that not just plant height but the weight of the corn grain produced was dramatically higher in the offspring. Agricultural scientists took note, and Shull's paper had a lasting impact on the field of applied genetics (**Figure 12.15**).

Farmers have known for centuries that matings among close relatives (known as **inbreeding**) can result in offspring of lower quality than matings between unrelated individuals. Agricultural scientists call this *in-*



		Sperm			
		BE	Be	bE	be
Eggs	BE	Black BBEE	Black BBEe	Black BbEE	Black BbEe
	Be	Black BBEe	Yellow BBee	Black BbEe	Yellow Bbee
	bE	Black BbEE	Black BbEe	Brown bbEE	Brown bbEe
	be	Black BbEe	Yellow Bbee	Brown bbEe	Yellow bbee

12.14 Genes May Interact Epistatically Epistasis occurs when one gene alters the phenotypic effect of another gene. In Labrador retrievers, the Ee gene determines the expression of the Bb gene.

breeding depression. The problems with inbreeding arise because close relatives tend to have the same recessive alleles, some of which may be harmful. The “hybrid vigor” after crossing inbred lines is called **heterosis** (short for heterozygosity). The cultivation of hybrid corn spread rapidly in the United States and all over the world, quadrupling grain production. Unfortunately, as we saw in the opening story, this scientific advance was not universally adopted, and regions such as the Russian empire fell far behind in corn production. The practice of hybridization has spread to many other crops and animals used in agriculture. For example, beef cattle that are crossbred are larger and live longer than cattle bred within their own genetic strain.

The mechanism by which heterosis works is not known. A widely accepted hypothesis is overdominance, in which the heterozygous condition in certain important genes whose products interact is superior to the homozygous condition in either or both genes. Another hypothesis is that the homozygotes have alleles that inhibit growth, and these are less active or absent in the heterozygote.

The environment affects gene action

The phenotype of an individual does not result from its genotype alone. *Genotype and environment interact to determine the phenotype of an organism.* This is especially important to remember in the era of genome sequencing (see Chapter 17). When the sequence of the human genome was completed in 2003, it was hailed as the “book of life,” and public expectations of the benefits gained from this knowledge were (and are) high. But this kind of “genetic determinism” is wrong. Common knowledge tells us that environmental variables such as light, temperature, and nutrition can affect the phenotypic expression of a genotype.



12.15 Hybrid Vigor in Corn Two homozygous parent lines of corn (cobs shown), B73 (left) and Mo17 (right), were crossed to produce the more vigorous hybrid line (center).

12.16 The Environment Influences Gene Expression This rabbit expresses a coat pattern known as “chocolate point.” Its genotype specifies dark fur, but the enzyme for dark fur is inactive at normal body temperature, so only the rabbit’s extremities—the coolest regions of the body—express this phenotype.

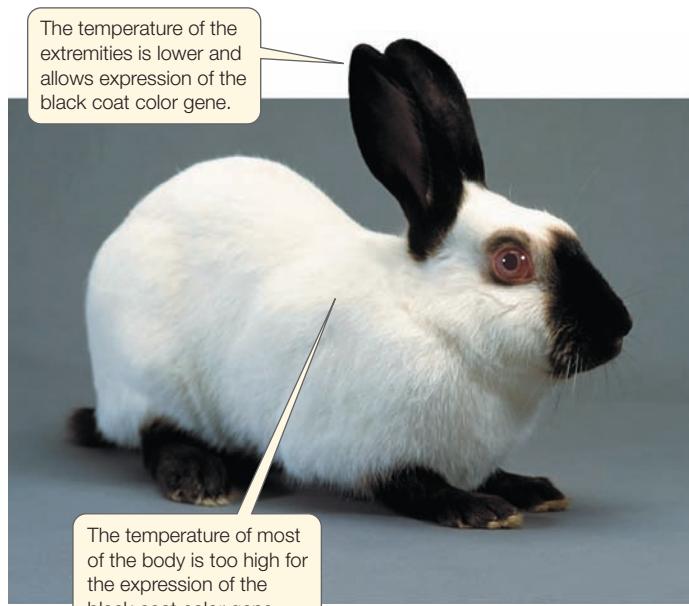
A familiar example of this phenomenon involves “point restriction” coat patterns found in Siamese cats and certain rabbit breeds (**Figure 12.16**). These animals carry a mutant allele of a gene that controls the growth of black fur all over the body. As a result of this mutation, the enzyme encoded by the gene is inactive at temperatures above a certain point (usually around 35°C). The animals maintain a body temperature above this point, and so their fur is mostly light. However, the extremities—feet, ears, nose, and tail—are cooler, about 25°C, so the fur on these regions is dark. These animals are all white when they are born, because the extremities were kept warm in the mother’s womb.

A simple experiment shows that the dark fur is temperature-dependent. If a patch of white fur on a point-restricted rabbit’s back is removed and an ice pack is placed on the skin where the patch was, the fur that grows back will be dark. This indicates that although the gene for dark fur was expressed all along, the environment inhibited the activity of the mutant enzyme.

Two parameters describe the effects of genes and environment on the phenotype:

- **Penetrance** is the proportion of individuals in a group with a given genotype that actually show the expected phenotype.
- **Expressivity** is the degree to which a genotype is expressed in an individual.

Penetrance affects, for example, the incidence of Huntington’s disease in humans. The disease results from the presence of a dominant allele, but 5 percent of people with the allele do not express the disease. So this allele is said to be 95 percent penetrant. For an example of environmental effects on expressivity, consider how Siamese cats kept indoors or outdoors in different climates might look.



Most complex phenotypes are determined by multiple genes and the environment

The differences between individual organisms in simple characters, such as those that Mendel studied in pea plants, are discrete and **qualitative**. For example, the individuals in a population of pea plants are either short or tall. For most complex characters, however, such as height in humans, the phenotype varies more or less continuously over a range. Some people are short, others are tall, and many are in between the two extremes. Such variation within a population is called **quantitative**, or continuous, variation (**Figure 12.17**).

12.17 Quantitative Variation Quantitative variation is produced by the interaction of genes at multiple loci and the environment. These students (women in white on the left are shorter; men in blue on the right are taller) show continuous variation in height that is the result of interactions between many genes and the environment.

Sometimes this variation is largely genetic. For instance, much of human eye color is the result of a number of genes controlling the synthesis and distribution of dark melanin pigment. Dark eyes have a lot of it, brown eyes less, and green, gray, and blue eyes even less. In the latter cases, the distribution of other pigments in the eye is what determines light reflection and color.

In most cases, however, quantitative variation is due to *both genes and environment*. Height in humans certainly falls into this category. If you look at families, you often see that parents and their offspring all tend to be tall or short. However, nutrition also plays a role in height: American 18-year-olds today are about 20 percent taller than their great-grandparents were at the same age. Three generations are not enough time for mutations that would exert such a dramatic effect to occur, so the height difference must not be due to genetics.

Geneticists call the genes that together determine such complex characters **quantitative trait loci**. Identifying these loci is a major challenge, and an important one. For example, the amount of grain that a variety of rice produces in a growing season is determined by many interacting genetic factors. Crop plant breeders have worked hard to decipher these factors in order to breed higher-yielding rice strains. In a similar way, human characteristics such as disease susceptibility and behavior are caused in part by quantitative trait loci. Recently, one of the many genes involved with human height was identified. The gene, *HMG A2*, has an allele that apparently has the potential to add 4 mm to human height.

12.3 RECAP

In epistasis, one gene affects the expression of another. Perhaps the most challenging problem for genetics is the explanation of complex phenotypes that are caused by many interacting genes and the environment.

- Explain the difference between penetrance and expressivity. **See p. 252**
- How is quantitative variation different from qualitative variation? **See pp. 252–253**

In the next section we'll see how the discovery that genes occupy specific positions on chromosomes enabled Mendel's successors to provide a physical explanation for his model of inheritance, and to provide an explanation for those cases where Mendel's second law does not apply.

12.4 What Is the Relationship between Genes and Chromosomes?

There are far more genes than chromosomes. Studies of different genes that are physically linked on the same chromosome reveal inheritance patterns that are not Mendelian. These patterns have been useful not only in detecting linkage of genes, but also in determining how far apart they are from one another on the chromosome.

The organism that revealed genetic linkage is the fruit fly *Drosophila melanogaster*. Its small size, the ease with which it can be bred, and its short generation time make this animal an attractive experimental subject. Beginning in 1909, Thomas Hunt Morgan and his students at Columbia University pioneered the study of *Drosophila*, and it remains a very important organism in studies of genetics.

yourBioPortal.com

GO TO Animated Tutorial 12.2 • Alleles That Do Not Sort Independently

Genes on the same chromosome are linked

Some of the crosses Morgan performed with fruit flies yielded phenotypic ratios that were not in accordance with those predicted by Mendel's law of independent assortment. Morgan crossed *Drosophila* with two known genotypes, *BbVvgvg* × *bbvgvg*,* for two different characters, body color and wing shape:

- *B* (wild-type gray body), is dominant over *b* (black body)
- *Vg* (wild-type wing) is dominant over *vg* (vestigial, a very small wing)

Morgan expected to see four phenotypes in a ratio of 1:1:1:1, but that is not what he observed. The body color gene and the wing size gene were not assorting independently; rather, they were, for the most part, inherited together (**Figure 12.18**).

These results became understandable to Morgan when he considered the possibility that the two loci are on the same chromosome—that is, that they might be linked. Suppose that the *B* and *Vg* loci are indeed located on the same chromosome. Why didn't all of Morgan's *F*₁ flies have the parental phenotypes—that is, why didn't his cross result in gray flies with normal wings (wild type) and black flies with vestigial wings, in a 1:1 ratio? If linkage were absolute—that is, if chromosomes always remained intact and unchanged—we would expect to see just those two types of progeny. However, this does not always happen.

Genes can be exchanged between chromatids

ABSOLUTE LINKAGE IS RARE If linkage were absolute, Mendel's law of independent assortment would apply only to loci on different chromosomes. What actually happens is more complex, and therefore more interesting. Genes at different loci on the same chromosome *do* sometimes separate from one another during meiosis. Genes may recombine when two homologous chromosomes physically exchange corresponding segments during prophase I of meiosis—that is, by crossing over (**Figure 12.19**; see also Figures 11.18 and 11.19). As described in Section 11.2, DNA is replicated during the S phase, so that by prophase I, when homologous chromosome pairs come together to form tetrads, each chromosome consists of two chromatids.

*Do you recognize this type of cross? It is a test cross for the two gene pairs; see Figure 12.6.

INVESTIGATING LIFE

12.18 Some Alleles Do Not Assort Independently

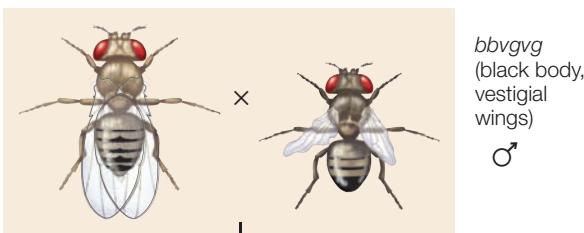
Morgan's studies showed that the genes for body color and wing size in *Drosophila* are linked, so that their alleles do not assort independently.

HYPOTHESIS Alleles for different characteristics always assort independently.

METHOD

Parent (P)

BbVvgv
Wild type
(gray body,
normal
wings)
♀



RESULTS

F₁

Genotypes

	<i>BbVvgv</i> Wild type	<i>bbvgvg</i> Black vestigial	<i>Bbvgvg</i> Gray vestigial	<i>bbVvgv</i> Black normal
--	----------------------------	----------------------------------	---------------------------------	-------------------------------

Expected phenotypes

575 575 575 575

Observed phenotypes
(number of individuals)

965	944	206	185
Parental phenotypes		Recombinant phenotypes	

These are the results expected from Mendel's second law (independent assortment)...

...but the actual results were inconsistent with the law.

the organism's parents). Usually several exchange events occur along the length of each homologous pair.

When crossing over takes place between two linked genes, not all the progeny of a cross have the parental phenotypes. Instead, recombinant offspring appear as well, as they did in Morgan's cross. They appear in proportions called **recombinant frequencies**, which are calculated by dividing the number of recombinant progeny by the total number of progeny (Figure 12.20). Recombinant frequencies will be greater for loci that are farther apart on the chromosome than for loci that are closer together because an exchange event is more likely to occur between genes that are far apart. Genetic recombination is another way to generate the diversity that is the raw material for natural selection and evolution.

CONCLUSION

The hypothesis is rejected. These two genes do not assort independently, but are linked (on the same chromosome).

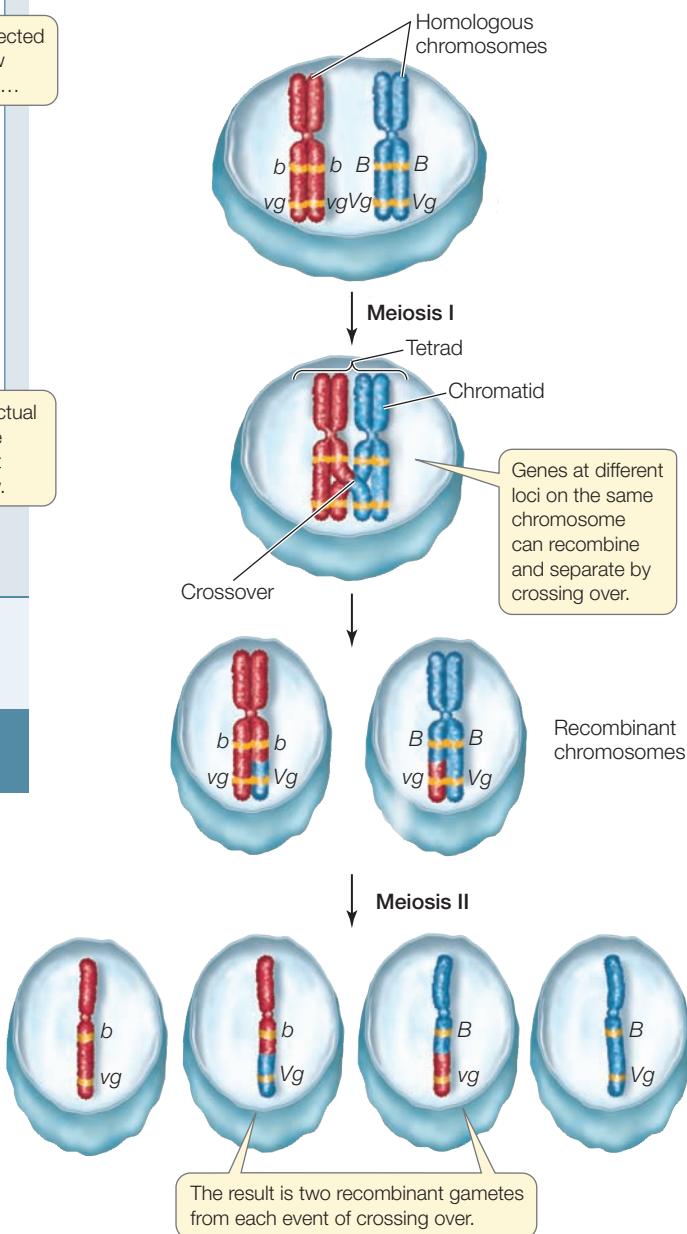
FURTHER INVESTIGATIONS: Look again at Mendel's dihybrid cross (see Figure 12.7). If the genes for seed shape and seed color were linked, what would the results be?

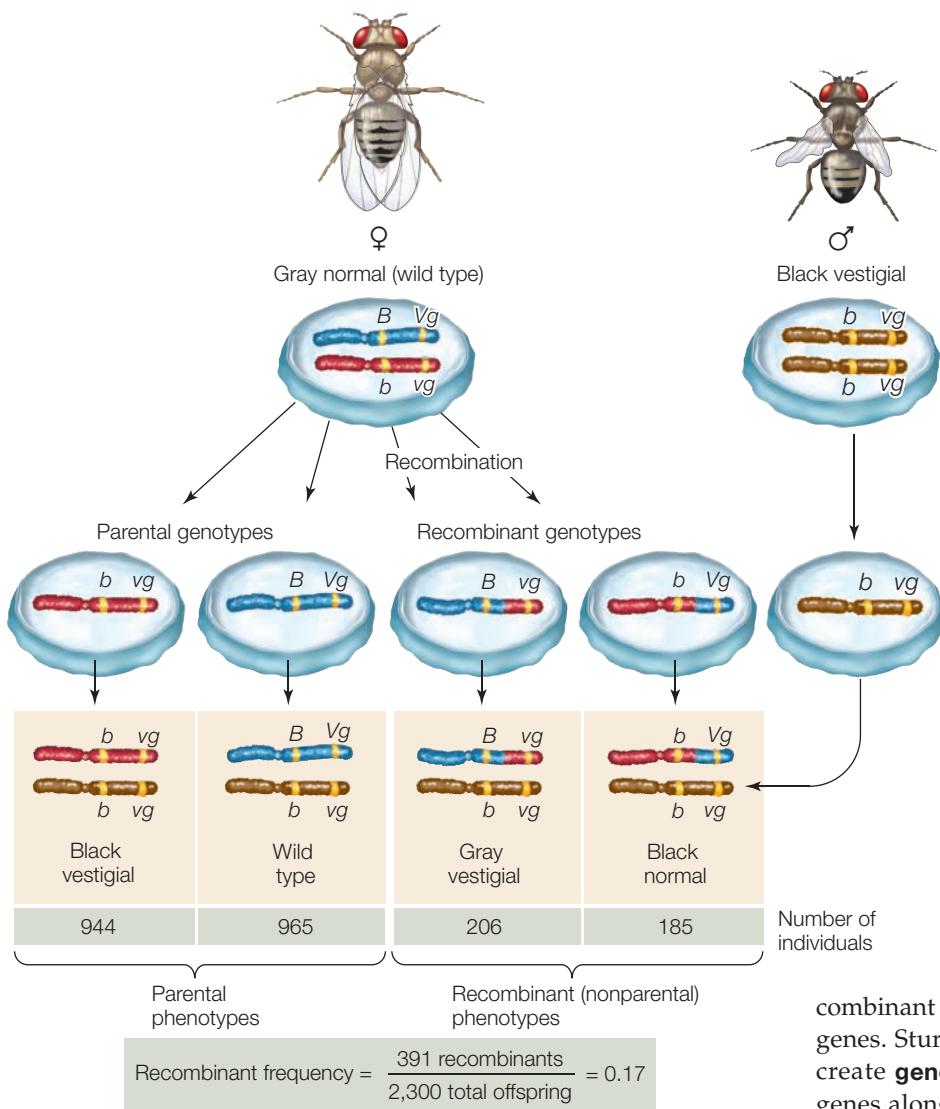
Go to yourBioPortal.com for original citations, discussions, and relevant links for all INVESTIGATING LIFE figures.

Note that the exchange event involves *only two of the four chromatids* in a tetrad, one from each member of the homologous pair, and can occur at any point along the length of the chromosome. The chromosome segments involved are exchanged reciprocally, so both chromatids involved in crossing over become recombinant (that is, each chromatid ends up with genes from both of

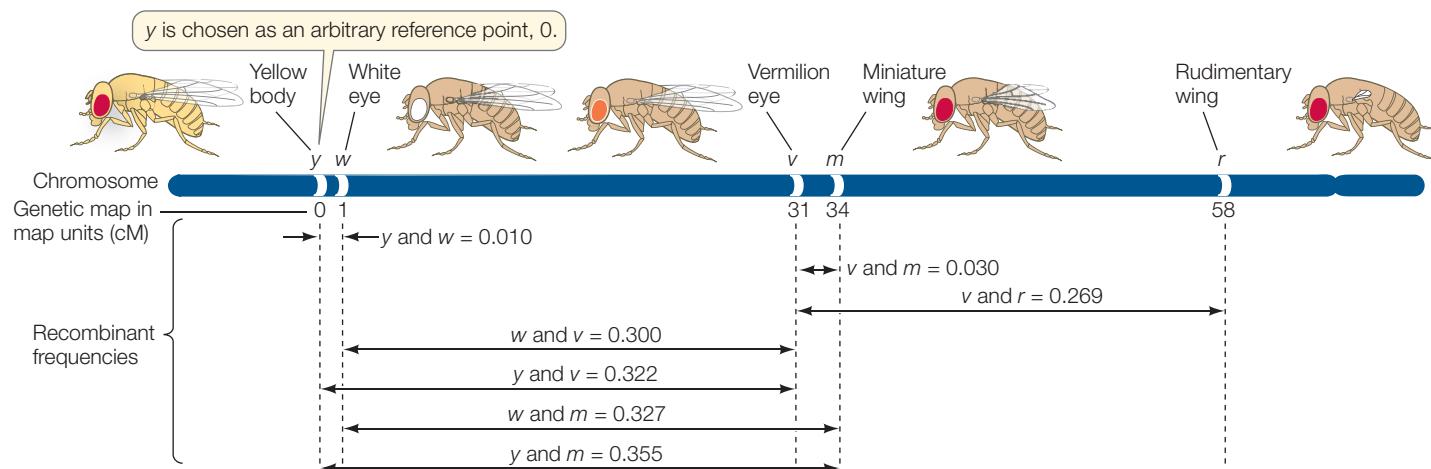
12.19 Crossing Over Results in Genetic Recombination

Recombination accounts for why linked alleles are not always inherited together. Alleles at different loci on the same chromosome can be recombined by crossing over, and separated from one another. Such recombination occurs during prophase I of meiosis.





12.21 Steps toward a Genetic Map The chance of a crossing over between two loci on a chromosome increases with the distance between the loci. Thus, Sturtevant was able to derive this partial map of a *Drosophila* chromosome using the Morgan group's data on the recombinant frequencies of five recessive traits. He used an arbitrary unit of distance—the map unit, or centimorgan (cM)—equivalent to a recombinant frequency of 0.01.



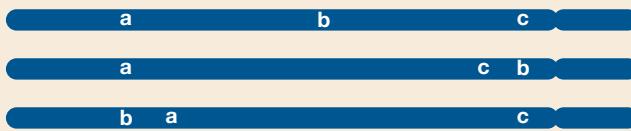
12.20 Recombinant Frequencies The frequency of recombinant offspring (those with a phenotype different from either parent) can be calculated.

Geneticists can make maps of chromosomes

If two loci are very close together on a chromosome, the odds of a crossover occurring between them are small. In contrast, if two loci are far apart, crossing over could occur between them at many points. This pattern is a consequence of the mechanism of crossing over: the farther apart two genes are, the more places there are in the chromosome for breakage and reunion of chromatids to occur. In a population of cells undergoing meiosis, a greater proportion of the cells will undergo recombination between two loci that are far apart than between two loci that are close together. In 1911, Alfred Sturtevant, then an undergraduate student in T. H. Morgan's fly room, realized how this simple insight could be used to show where different genes lie on a chromosome in relation to one another.

The Morgan group had determined recombinant frequencies for many pairs of linked *Drosophila* genes. Sturtevant used those recombinant frequencies to create **genetic maps** that showed the arrangements of genes along the chromosomes (Figure 12.21). Ever since Sturtevant demonstrated this method, geneticists have mapped the chromosomes of eukaryotes, prokaryotes, and viruses, assigning distances between genes in **map units**. A map unit corresponds to a recombinant frequency of 0.01; it is also referred to as a **centimorgan (cM)**, in honor of the founder of the fly room. You, too, can work out a genetic map (Figure 12.22).

- 1 At the outset, we have no idea of the individual distances between the genes, and there are several possible sequences (a-b-c, a-c-b, b-a-c).



We make a cross $AABB \times aabb$, and obtain an F_1 generation with a genotype $AaBb$. We test cross these $AaBb$ individuals with $aabb$. Here are the genotypes of the first 1,000 progeny:

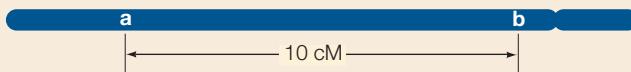
450 $AaBb$, 450 $aabb$, 50 $Aabb$, and 50 $aaBb$.
(parental types) (recombinant types)

2 How far apart are the a and b genes?

What is the recombinant frequency? Which are the recombinant types, and which are the parental types?

Recombinant frequency (a to b) = $(50 + 50)/1,000 = 0.1$
So the map distance is

Map distance = $100 \times \text{recombinant frequency} = 100 \times 0.1 = 10 \text{ cM}$



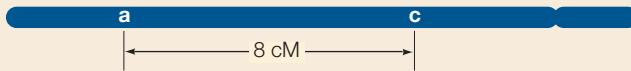
3 How far apart are the a and c genes?

Now we make a cross $AACC \times aacc$, obtain an F_1 generation, and test cross it, obtaining

460 $AaCc$, 460 $aacc$, 40 $Aacc$, and 40 $aaCc$

Recombinant frequency (a to c) = $(40 + 40)/1,000 = 0.08$

Map distance = $100 \times \text{recombinant frequency} = 100 \times 0.08 = 8 \text{ cM}$



Linkage is revealed by studies of the sex chromosomes

In Mendel's work, reciprocal crosses always gave identical results; it did not matter whether a dominant allele was contributed by the mother or by the father. But in some cases, the parental origin of a chromosome does matter. For example, human males inherit a bleeding disorder called hemophilia from their mothers, not from their fathers. To understand the types of inheritance in which the parental origin of an allele is important, we must consider the ways in which sex is determined in different species.

SEX DETERMINATION BY CHROMOSOMES In corn, every diploid adult has both male and female reproductive structures. The tissues in these two types of structure are genetically identical, just as roots and leaves are genetically identical. Plants such as corn, in which the same individual produces both male and female gametes, are said to be *monoecious* (Greek, "one house"). Other plants, such as date palms and oak trees, and most animals are *dioecious* ("two houses"), meaning that some individuals can produce only male gametes and the others can produce only fe-

12.22 Map These Genes The object of this exercise is to determine the order of three loci (a, b, and c) on a chromosome, as well as the map distances (in cM) between them.

4 How far apart are the b and c genes?

We make a cross $BBCC \times bbcc$, obtain an F_1 generation, and test cross it, obtaining

490 $BbCc$, 490 $bbcc$, 10 $Bbcc$, and 10 $bbCc$

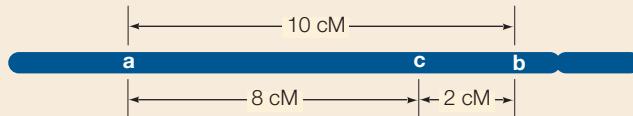
Recombinant frequency (b to c) = $(10 + 10)/1,000 = 0.02$

Map distance = $100 \times \text{recombinant frequency} = 100 \times 0.02 = 2 \text{ cM}$



5 Which of the three genes is between the other two?

Because a and b are the farthest apart, c must be between them.



These numbers add up perfectly. In most real cases, they will not add up perfectly because of multiple crossovers.

male gametes. In other words, in dioecious organisms the different sexes are different individuals.

In most dioecious organisms, sex is determined by differences in the chromosomes, but such determination operates in different ways in different groups of organisms. For example, in many animals including mammals, sex is determined by a single **sex chromosome**, or by a pair of them. Both males and females have two copies of each of the rest of the chromosomes, which are called **autosomes**. In other animals, the chromosomal basis of sex determination is different from that of mammals (**Table 12.2**).

The sex chromosomes of female mammals consist of a pair of X chromosomes. Male mammals, on the other hand, have one X chromosome and a sex chromosome that is not found in fe-

TABLE 12.2
Sex Determination in Animals

ANIMAL GROUP	MECHANISM
Bees	Males are diploid, females are haploid
Fruit Flies	Fly is female if ratio of sex chromosomes to autosomes is ≥ 1
Birds	Males WW (homogametic), females WZ (heterogametic)
Mammals	Males XY (heterogametic), females XX (homogametic)

males, the Y chromosome. Females may be represented as XX and males as XY.

MALE MAMMALS PRODUCE TWO KINDS OF GAMETES Each gamete produced by a male mammal has a complete set of autosomes, but half the gametes carry an X chromosome, and the other half carry a Y. When an X-bearing sperm fertilizes an egg, the resulting XX zygote is female; when a Y-bearing sperm fertilizes an egg, the resulting XY zygote is male.

SEX CHROMOSOME ABNORMALITIES REVEALED THE GENE THAT DETERMINES SEX Can we determine which chromosome, X or Y, carries the sex-determining gene, and can the gene be identified? One way to determine cause (e.g., the presence of a gene on the Y chromosome) and effect (e.g., maleness) is to look at cases of biological error, in which the expected outcome does not happen.

Abnormal sex chromosome arrangements resulting from nondisjunction during meiosis (see Section 11.5) tell us something about the functions of the X and Y chromosomes. As you will recall, nondisjunction occurs when a pair of homologous chromosomes (in meiosis I) or sister chromatids (in meiosis II) fail to separate. As a result, a gamete may have one too few or one too many chromosomes. If this gamete fuses with another gamete that has the full haploid chromosome set, the resulting offspring will be aneuploid, with fewer or more chromosomes than normal.

In humans, XO individuals sometimes appear. (The O implies that a chromosome is missing—that is, individuals that are XO have only one sex chromosome.) Human XO individuals are females who are moderately abnormal physically but normal mentally; usually they are also sterile. The XO condition in humans is called Turner syndrome. It is the only known case in which a person can survive with only one member of a chromosome pair (here, the XY pair), although most XO conceptions are spontaneously terminated early in development. XXY individuals also occur; this condition, which affects males, is called Klinefelter syndrome, and results in overlong limbs and sterility.

These observations suggest that the gene controlling maleness is located on the Y chromosome. Observations of people with other types of chromosomal abnormalities helped researchers to pinpoint the location of that gene:

- Some women are genetically XY but lack a small portion of the Y chromosome.
- Some men are genetically XX but have a small piece of the Y chromosome attached to another chromosome.

The Y fragments that are respectively missing and present in these two cases are the same and contain the maleness-determining gene, which was named *SRY* (sex-determining region on the Y chromosome).

The *SRY* gene encodes a protein involved in **primary sex determination**—that is, the determination of the kinds of gametes that an individual will produce and the organs that will make them. In the presence of the functional *SRY* protein, an embryo develops sperm-producing testes. (Notice that *italic type* is used for the name of a gene, but *roman type* is used for the

name of a protein.) If the embryo has no Y chromosome, the *SRY* gene is absent, and thus the *SRY* protein is not made. In the absence of the *SRY* protein, the embryo develops egg-producing ovaries. In this case, a gene on the X chromosome called *DAX1* produces an anti-testis factor. So the role of *SRY* in a male is to inhibit the maleness inhibitor encoded by *DAX1*. The *SRY* protein does this in male cells, but since it is not present in females, *DAX1* can act to inhibit maleness.

Primary sex determination is not the same as **secondary sex determination**, which results in the outward manifestations of maleness and femaleness (such as body type, breast development, body hair, and voice). These outward characteristics are not determined directly by the presence or absence of the Y chromosome. Instead, they are determined by genes that are scattered on the autosomes and the X chromosome. These genes control the actions of hormones, such as testosterone and estrogen.

Genes on sex chromosomes are inherited in special ways

Genes on sex chromosomes do not show the Mendelian patterns of inheritance. In *Drosophila* and in humans, the Y chromosome carries few known genes, but the X chromosome carries a substantial number of genes that affect a great variety of characters. These genes are present in two copies in females but only one copy in males. Therefore, males are always **hemizygous** for genes on the X chromosome—they have only one copy of each, and it is expressed. So reciprocal crosses do not give identical results for characters whose genes are carried on the sex chromosomes, and these characters do not show the usual Mendelian inheritance ratios.

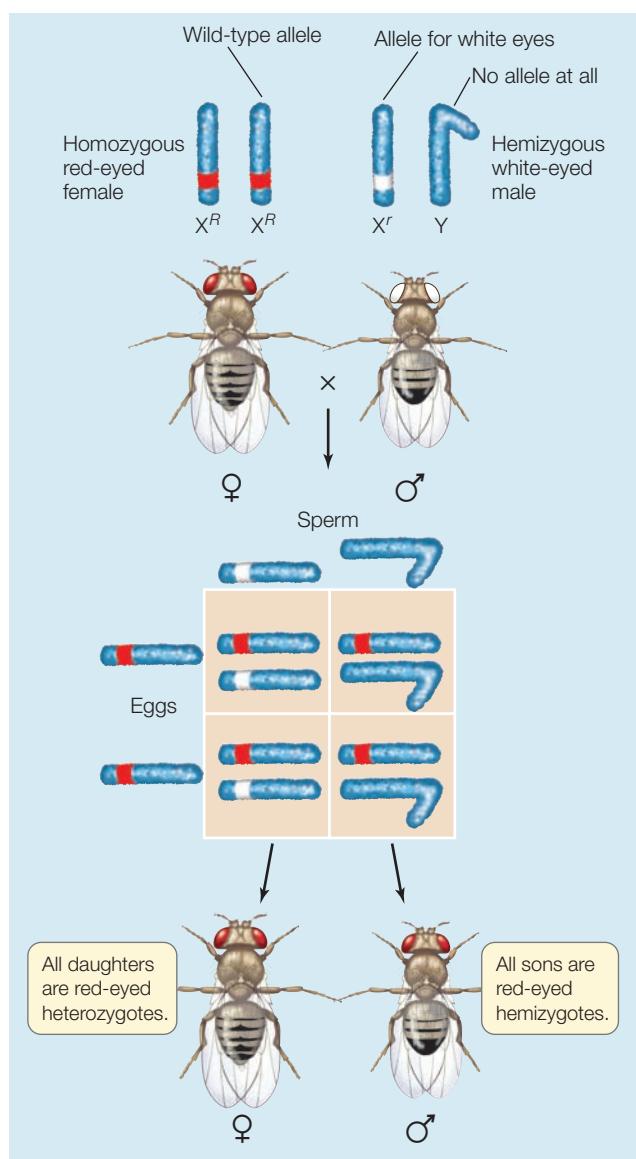
Eye color in *Drosophila* is a good example of inheritance of a character that is governed by a locus on a sex chromosome (**sex-linked inheritance**). The wild-type eye color of these flies is red. In 1910, Morgan discovered a mutation that causes white eyes. He crossed flies of the wild-type and mutant phenotypes, and demonstrated that the eye color locus is on the X chromosome. If we abbreviate the eye color alleles as *R* (red eyes) and *r* (white eyes), the presence of the alleles on the X chromosome is designated by X^R and X^r .

When a homozygous red-eyed female ($X^R X^R$) was crossed with a (hemizygous) white-eyed male ($X^r Y$), all the sons and daughters had red eyes, because red (*R*) is dominant over white (*r*) and all the progeny had inherited a wild-type X chromosome (X^R) from their mothers (**Figure 12.23A**).

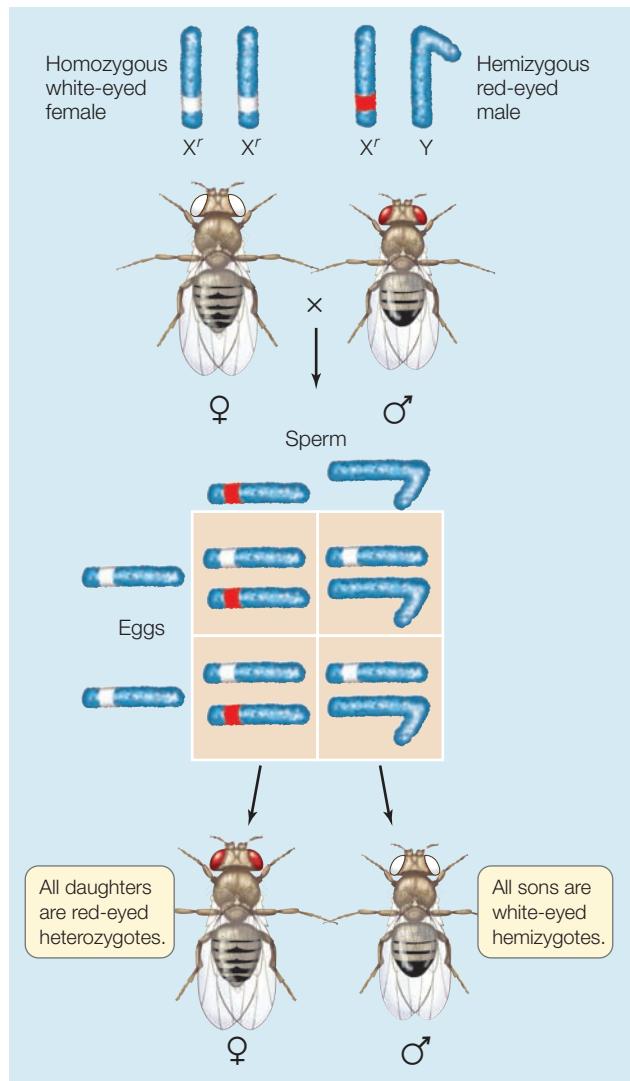
In the reciprocal cross, in which a white-eyed female ($X^r X^r$) was mated with a red-eyed male ($X^R Y$), all the sons were white-eyed and all the daughters were red-eyed (**Figure 12.23B**). The sons from the reciprocal cross inherited their only X chromosome from their white-eyed mother; the Y chromosome they inherited from their father did not carry the eye color locus. On the other hand, the daughters got an X chromosome bearing the white allele from their mother and an X chromosome bearing the red allele from their father; therefore they were red-eyed heterozygotes.

When heterozygous females were mated with red-eyed males, half their sons had white eyes, but all their daughters

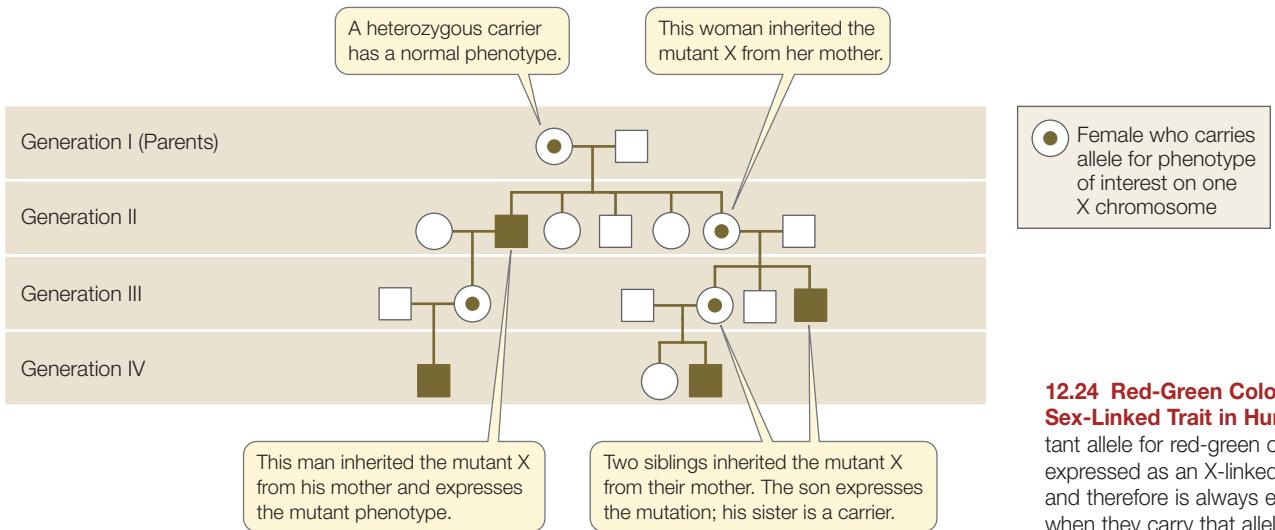
(A)



(B)



12.23 Eye Color Is a Sex-Linked Trait in *Drosophila* Morgan demonstrated that a mutant allele that causes white eyes in *Drosophila* is carried on the X chromosome. Note that in this case, the reciprocal crosses do not have the same results.



12.24 Red-Green Color Blindness Is a Sex-Linked Trait in Humans The mutant allele for red-green color blindness is expressed as an X-linked recessive trait, and therefore is always expressed in males when they carry that allele.

had red eyes. Together, these three results showed that eye color was carried on the X chromosome and not on the Y.

Humans display many sex-linked characters

The human X chromosome carries about 2,000 known genes. The alleles at these loci follow the same pattern of inheritance as those for eye color in *Drosophila*. For example, one gene on the human X chromosome has a mutant recessive allele that leads to red-green color blindness, and it appears in individuals who are homozygous or hemizygous for the recessive mutant allele.

Pedigree analyses of X-linked recessive phenotypes (like the one in **Figure 12.24**) reveal the following patterns:

- The phenotype appears much more often in males than in females, because only one copy of the rare allele is needed for its expression in males, while two copies must be present in females.
- A male with the mutation can pass it on only to his daughters; all his sons get his Y chromosome.
- Daughters who receive one X-linked mutation are heterozygous **carriers**. They are phenotypically normal, but they can pass the mutant allele to either sons or daughters. (On average only half their children inherit the mutant allele, since half of their X chromosomes carry the normal allele.)
- The mutant phenotype can skip a generation if the mutation passes from a male to his daughter (who will be phenotypically normal) and thus to her son.

The small human Y chromosome carries several dozen genes. Among them is the maleness determinant, *SRY*. Interestingly, for some genes on the Y chromosome there are similar, but not identical, genes on the X chromosome. For example, one of the proteins that make up ribosomes is encoded by a gene on the Y chromosome that is expressed only in male cells, while the X-linked counterpart is expressed in both sexes. This means that there are “male” and “female” ribosomes; the significance of this phenomenon is unknown. Y-linked alleles are passed only from father to son. (Verify this with a Punnett square.)

12.4 RECAP

Simple Mendelian ratios are not observed when genes are linked on the same chromosome. Linkage is indicated by atypical frequencies of phenotypes in the offspring from a test cross. Sex linkage in humans refers to genes on the X chromosome that have no counterpart on the Y chromosome.

- What is the concept of linkage and what are its implications for the results of genetic crosses? **See pp. 253–254 and Figures 12.19 and 12.20**
- How does a sex-linked gene behave differently in genetic crosses than a gene on an autosome? **See pp. 257–259 and Figure 12.23**

The genes we’ve discussed so far in this chapter are all in the cell nucleus. But other organelles, including mitochondria and plastids, also carry genes. What are they, and how are they inherited?

12.5 What Are the Effects of Genes Outside the Nucleus?

The nucleus is not the only organelle in a eukaryotic cell that carries genetic material. As described in Section 5.5, mitochondria and plastids contain small numbers of genes, which are remnants of the entire genomes of colonizing prokaryotes that eventually gave rise to these organelles. For example, in humans, there are about 24,000 genes in the nuclear genome and 37 in the mitochondrial genome. Plastid genomes are about five times larger than those of mitochondria. In any case, several of the genes carried by cytoplasmic organelles are important for organelle assembly and function, so it is not surprising that mutations of these genes can have profound effects on the organism.

The inheritance of organelle genes differs from that of nuclear genes for several reasons:

- In most organisms, mitochondria and plastids are inherited only from the mother. As you will learn in Chapter 43, eggs contain abundant cytoplasm and organelles, but the only part of the sperm that survives to take part in the union of haploid gametes is the nucleus. So you have inherited your mother’s mitochondria (with their genes), but not your father’s.
- There may be hundreds of mitochondria or plastids in a cell. So a cell is not diploid for organelle genes.
- Organelle genes tend to mutate at much faster rates than nuclear genes, so there are multiple alleles of organelle genes.

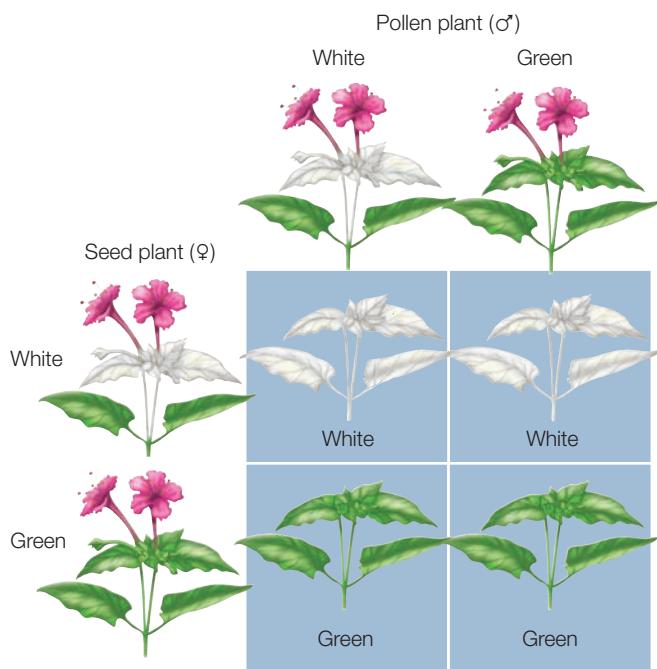
The phenotypes resulting from mutations in organelle genes reflect the organelles’ roles. For example, in plants and some photosynthetic protists, certain plastid gene mutations affect the proteins that assemble chlorophyll molecules into photosystems. These mutations result in a phenotype that is essentially white instead of green. The inheritance of this phenotype follows a non-Mendelian, maternal pattern (**Figure 12.25**).

Mitochondrial gene mutations that affect one of the complexes in the respiratory chain result in less ATP production. These mutations have particularly noticeable effects in tissues with high energy requirements, such as the nervous system, muscles, and kidneys. In 1995, Greg LeMond, a professional cyclist who had won the famous Tour de France three times, was forced to retire because of muscle weakness caused by a mitochondrial mutation.

12.5 RECAP

Genes in the genomes of organelles, specifically plastids and mitochondria, do not behave in a Mendelian fashion.

- Why are genes carried in the organelle genomes usually inherited only from the mother?



12.25 Cytoplasmic Inheritance In four o'clock plants, leaf color is inherited through the female plant only. The white leaf color is caused by a chloroplast mutation that occurs during the life of the parent plant; the leaves that form before the mutation occurs are green. The mutation is passed on to the germ cells, and the offspring that inherit the mutation are entirely white.

Mendel and those who followed him scientifically focused on eukaryotes, with diploid organisms and haploid gametes. A half-century after the rediscovery of Mendel's work, a sexual process that allows genetic recombination was discovered in prokaryotes as well. We now turn to that process.

12.6 How Do Prokaryotes Transmit Genes?

As you saw in Chapter 5, prokaryotic cells lack a nucleus but contain their genetic material as mostly a single chromosome in a central region of the cell. In Chapter 11, you saw that bacteria reproduce asexually by cell division, a process that gives rise to virtually genetically identical products. That is, the offspring of cell reproduction in bacteria constitute a clone. However, mutations occur in bacteria just as they do in eukaryotes; the resulting new alleles increase genetic diversity.

You might expect, therefore, that there is no way for individuals of these species to exchange genes, as in sexual reproduction. It turns out, though, that prokaryotes do have a sexual process.

Bacteria exchange genes by conjugation

The bacterial chromosome, like the bacterial cell, is considerably smaller than its eukaryotic counterpart. In humans, each of the 23 chromosomes in a haploid set may have thousands of linked genes and be a highly compacted linear strand several centimeters in length. In contrast, *E. coli* has a single, circular chromosome that carries a few thousand genes and is only about 1 μm in circumference. Genetic recombination in bacte-

ria occurs after a chromosome is transferred from one cell to another, which brings the chromosomes of two cells into close proximity within a single cell.

Joshua Lederberg and Edward Tatum discovered this recombination process in 1946. They worked with two genetic strains of *E. coli* that had different alleles for each of six genes (each of the genes coded for the synthesis of certain small molecules). Simply put, the two strains had the following genotypes (remember that bacteria are haploid):

$$ABCdef \quad \text{and} \quad abcDEF$$

where capital letters stand for wild-type alleles and lower case letters stand for mutant alleles.

When the two strains were grown in the same environment in the laboratory, most of the cells produced clones. That is, almost all of the cells that grew had the original genotypes:

$$ABCdef \quad \text{and} \quad abcDEF$$

However, very rarely, Lederberg and Tatum detected bacteria that had the genotype

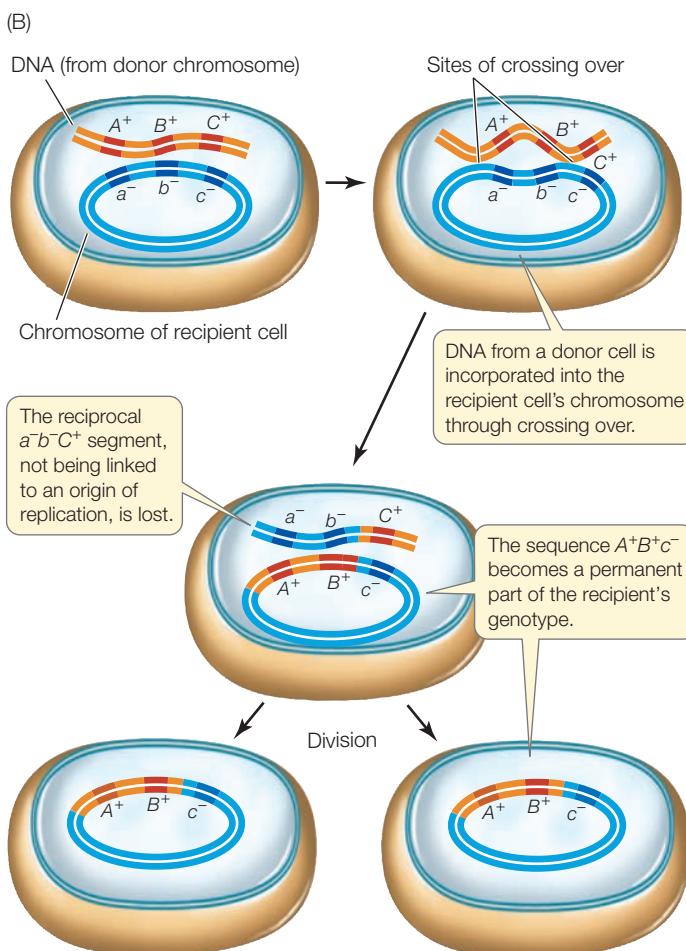
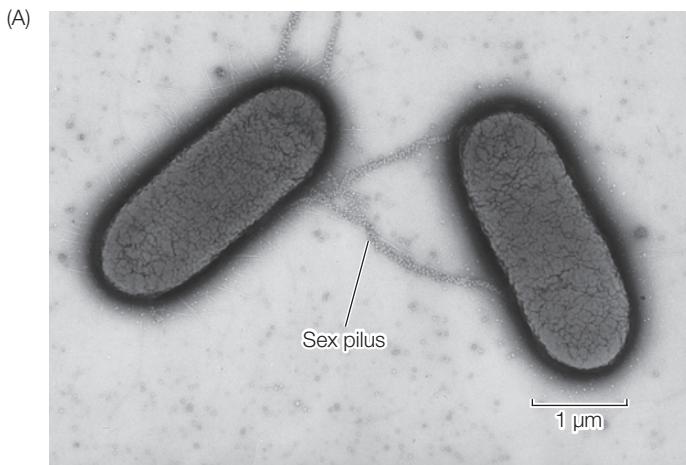
$$ABCDEF$$

How could these completely wild-type bacteria have arisen? One possibility was mutation: the *d* allele could have mutated to *D*, and so on for *e* and *f*. The problem with this explanation was that the probabilities of mutation from *d* to *D*, *e* to *E*, and *f* to *F* were each very low. The probability of all three events occurring in the same cell would be the product of the three individual probabilities—an extremely low number and millions of times lower than the actual rate of appearance of cells with the genotype *ABCDEF*.

Electron microscopy showed how sexual transmission in bacteria might happen, via physical contact between the cells (Figure 12.26A). Physical contact is initiated by a thin projection called a **sex pilus** (plural *pili*). Once sex pili bring the two cells together, the actual transfer of the chromosome occurs through a thin cytoplasmic bridge called a **conjugation tube** that forms between the cells.

The chromosome moves in a linear fashion from a donor cell to a recipient cell. Since the bacterial chromosome is circular, it must be made linear (cut) before it can pass through the tube. Contact between the cells is brief—only rarely long enough for the entire donor genome to enter the recipient cell. Therefore, the recipient cell usually receives only a portion of the donor chromosome. There is no reciprocal transfer of a chromosome from the recipient to the donor.

Once the donor chromosome fragment is inside the recipient cell, it can recombine with the recipient cell's chromosome. In much the same way that chromosomes pair up, gene for gene, in prophase I of meiosis, the donor chromosome can line up beside its homologous genes in the recipient, and crossing over can occur. Gene(s) from the donor can become integrated into the genome of the recipient, thus changing the recipient's genetic constitution (Figure 12.26B), although only about half the transferred genes become integrated in this way. When the recipient cells proliferate, the donor genes are passed on to all progeny cells.



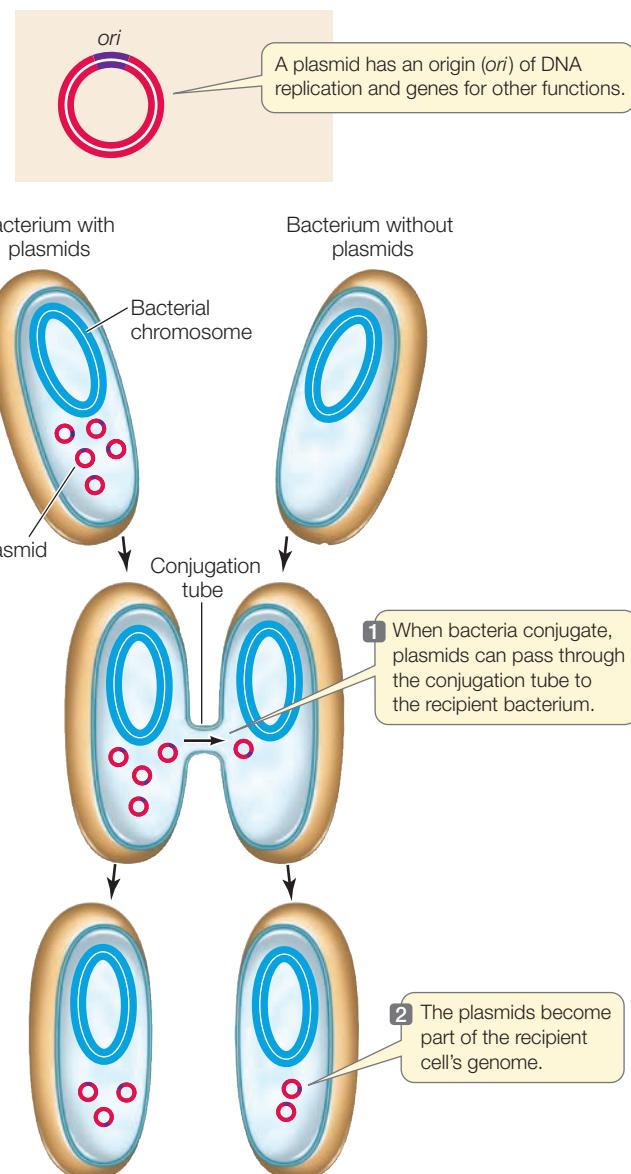
12.26 Bacterial Conjugation and Recombination (A) Sex pili draw two bacteria into close contact, so that a cytoplasmic conjugation tube can form. DNA is transferred from one cell to the other via the conjugation tube. (B) DNA from a donor cell can become incorporated into a recipient cell's chromosome through crossing over.

Plasmids transfer genes between bacteria

In addition to their main chromosome, many bacteria harbor additional smaller, circular chromosomes called **plasmids**. They typically contain at most a few dozen genes, which, depending on the particular plasmid, may fall into one of several categories:

- Genes for unusual metabolic capacities, such as the ability to break down hydrocarbons; bacteria carrying these plasmids can be used to clean up oil spills.
- Genes for conjugation, including the ability to make a sex pilus; bacteria carrying this type of plasmid, called fertility factor, are designated F^+ and conjugate with bacteria that lack the plasmid (F^-).
- Genes for antibiotic resistance; bacteria carrying such gene(s)—the plasmids are called R factors—are a major threat to human health.

Plasmids can move between cells during conjugation, thereby transferring new genes to the recipient bacterium (Figure 12.27). Because plasmids can replicate independently of the main chromosome, they do not need to recombine with the main chromosome to add their genes to the recipient cell's genome.



12.27 Gene Transfer by Plasmids When plasmids enter a cell via conjugation, their genes can be expressed in the recipient cell.

12.6 RECAP

Although they are haploid and reproduce primarily asexually, prokaryotes have the ability to transfer genes from one cell to another. These genes can be part of the main single chromosome or on a small chromosome called a plasmid.

- How were prokaryotic gene transfer and recombination discovered? [See p. 260](#)
- What are the differences between recombination after conjugation in prokaryotes and recombination during meiosis in eukaryotes?

CHAPTER SUMMARY

12.1 What Are the Mendelian Laws of Inheritance?

- Physical features of organisms, or **characters**, can exist in different forms, or **traits**. A **heritable trait** is one that can be passed from parent to offspring. A **phenotype** is the physical appearance of an organism; a **genotype** is the genetic constitution of the organism.
- The different forms of a **gene** are called **alleles**. Organisms that have two identical alleles for a trait are called **homozygous**; organisms that have two different alleles for a trait are called **heterozygous**. A gene resides at a particular site on a chromosome called a **locus**.
- Mendel's experiments included **reciprocal crosses** and **mono-hybrid crosses** between **true-breeding** pea plants. Analysis of his meticulously tabulated data led Mendel to propose a **particulate theory** of inheritance stating that discrete units (now called genes) are responsible for the inheritance of specific traits, to which both parents contribute equally.
- Mendel's first law, the **law of segregation**, states that when any individual produces gametes, the two copies of a gene separate, so that each gamete receives only one member of the pair. Thus every individual in the F₁ inherits one copy from each parent. [Review Figures 12.4 and 12.5](#)
- Mendel used a **test cross** to find out whether an individual showing a dominant phenotype was homozygous or heterozygous. [Review Figure 12.6, WEB ACTIVITY 12.1](#)
- Mendel's use of **dihybrid crosses** to study the inheritance of two characters led to his second law: the **law of independent assortment**. The independent assortment of genes in meiosis leads to **recombinant** phenotypes. [Review Figures 12.7 and 12.8, ANIMATED TUTORIAL 12.1](#)
- Probability calculations and **pedigrees** help geneticists trace Mendelian inheritance patterns. [Review Figures 12.9 and 12.10](#)

12.2 How Do Alleles Interact?

- New alleles arise by random **mutation**. Many genes have multiple alleles. A **wild-type** allele gives rise to the predominant form of a trait. When the wild-type allele is present at a locus less than 99 percent of the time, the locus is said to be **polymorphic**. [Review Figure 12.11](#)
- In **incomplete dominance**, neither of two alleles is dominant. The heterozygous phenotype is intermediate between the homozygous phenotypes. [Review Figure 12.12](#)

- Codominance** exists when two alleles at a locus produce two different phenotypes that both appear in heterozygotes.
- An allele that affects more than one trait is said to be **pleiotropic**.

12.3 How Do Genes Interact?

- In **epistasis**, one gene affects the expression of another. [Review Figure 12.14](#)
- Environmental conditions can affect the expression of a genotype.
- Penetrance** is the proportion of individuals in a group with a given genotype that show the expected phenotype. **Expressivity** is the degree to which a genotype is expressed in an individual.
- Variations in phenotypes can be **qualitative** (discrete) or **quantitative** (graduated, continuous). Most quantitative traits are the result of the effects of several genes and the environment. Genes that together determine quantitative characters are called **quantitative trait loci**.

12.4 What Is the Relationship between Genes and Chromosomes?

SEE ANIMATED TUTORIAL 12.2

- Each chromosome carries many genes. Genes on the same chromosome are referred to as a **linkage group**.
- Genes on the same chromosome can recombine by crossing over. The resulting recombinant chromosomes have new combinations of alleles. [Review Figures 12.19 and 12.20](#)
- Sex chromosomes** carry genes that determine whether the organism will produce male or female gametes. All other chromosomes are called **autosomes**. The specific functions of X and Y chromosomes differ among different groups of organisms.
- Primary sex determination** in mammals is usually a function of the presence or absence of the **SRY** gene. **Secondary sex determination** results in the outward manifestations of maleness or femaleness.
- In fruit flies and mammals, the X chromosome carries many genes, but the Y chromosome has only a few. Males have only one allele (are **hemizygous**) for X-linked genes, so recessive **sex-linked** mutations are expressed phenotypically more often in males than in females. Females may be unaffected **carriers** of such alleles.

12.5 What Are the Effects of Genes Outside the Nucleus?

- Cytoplasmic organelles such as plastids and mitochondria contain small numbers of genes. In many organisms, cytoplasmic genes are inherited only from the mother because the male gamete contributes only its nucleus (i.e., no cytoplasm) to the zygote at fertilization. **Review Figure 12.25**

12.6 How Do Prokaryotes Transmit genes?

- Prokaryotes reproduce primarily asexually but can exchange genes in a sexual process called conjugation. **Review Figure 12.26**
- Plasmids are small, extra chromosomes in bacteria that carry genes involved in important metabolic processes and that can be transmitted from one cell to another. **Review Figure 12.27**

SEE WEB ACTIVITIES 12.2 and 12.3 for a concept review of this chapter.

SELF-QUIZ

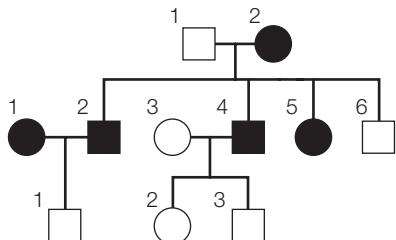
- In a simple Mendelian monohybrid cross, true-breeding tall plants are crossed with short plants, and the F_1 plants, which are all tall, are allowed to self-pollinate. What fraction of the F_2 generation are both tall and heterozygous?
 - 1/8
 - 1/4
 - 1/3
 - 2/3
 - 1/2
- The phenotype of an individual
 - depends at least in part on the genotype.
 - is either homozygous or heterozygous.
 - determines the genotype.
 - is the genetic constitution of the organism.
 - is either monohybrid or dihybrid.
- The ABO blood groups in humans are determined by a multiple-allele system in which I^A and I^B are codominant and are both dominant to I^O . A newborn infant is type A. The mother is type O. Possible phenotypes of the father are
 - A, B, or AB.
 - A, B, or O.
 - O only.
 - A or AB.
 - A or O.
- Which statement about an individual that is homozygous for an allele is *not* true?
 - Each of its cells possesses two copies of that allele.
 - Each of its gametes contains one copy of that allele.
 - It is true-breeding with respect to that allele.
 - Its parents were necessarily homozygous for that allele.
 - It can pass that allele to its offspring.
- Which statement about a test cross is *not* true?
 - It tests whether an unknown individual is homozygous or heterozygous.
 - The test individual is crossed with a homozygous recessive individual.
 - If the test individual is heterozygous, the progeny will have a 1:1 ratio.
 - If the test individual is homozygous, the progeny will have a 3:1 ratio.
 - Test cross results are consistent with Mendel's model of inheritance for unlinked genes.
- Linked genes
 - must be immediately adjacent to one another on a chromosome.
 - have alleles that assort independently of one another.
 - never show crossing over.
 - are on the same chromosome.
 - always have multiple alleles.
- In the F_2 generation of a dihybrid cross
 - four phenotypes appear in the ratio 9:3:3:1 if the loci are linked.
 - four phenotypes appear in the ratio 9:3:3:1 if the loci are unlinked.
 - two phenotypes appear in the ratio 3:1 if the loci are unlinked.
 - three phenotypes appear in the ratio 1:2:1 if the loci are unlinked.
 - two phenotypes appear in the ratio 1:1 whether or not the loci are linked.
- The genetic sex of a human is determined by
 - ploidy, with the male being haploid.
 - the Y chromosome.
 - X and Y chromosomes, the male being XX.
 - the number of X chromosomes, the male being XO.
 - Z and W chromosomes, the male being ZZ.
- In epistasis
 - nothing changes from generation to generation.
 - one gene alters the effect of another.
 - a portion of a chromosome is deleted.
 - a portion of a chromosome is inverted.
 - the behavior of two genes is entirely independent.
- In humans, spotted teeth are caused by a dominant sex-linked gene. A man with spotted teeth whose father had normal teeth marries a woman with normal teeth. Therefore,
 - all of their daughters will have normal teeth.
 - all of their daughters will have spotted teeth.
 - all of their children will have spotted teeth.
 - half of their sons will have spotted teeth.
 - all of their sons will have spotted teeth.

GENETICS PROBLEMS

1. In guinea pigs, black body color (*B*) is completely dominant over albino (*b*). For the crosses below, give the genotypes of the parents:

Parental phenotypes	Black offspring	Albino offspring	Parental genotypes?
Black × albino	12	0	
Albino × albino	0	12	
Black × albino	5	7	
Black × black	9	3	

2. In the genetic cross, *AaBbCcDdEE* × *AaBBCcDdEe*, what fraction of the offspring will be heterozygous for all of these genes (*AaBbCcDdEe*)? Assume all genes are unlinked and the alleles show simple dominance.
3. The pedigree below shows the inheritance of a rare mutant phenotype in humans, congenital cataracts (black symbols).



- a. Are cataracts inherited as an autosomal dominant trait? Autosomal recessive? Sex-linked dominant? Sex-linked recessive?
- b. Person #5 in the second generation marries a man who does not have cataracts. Two of their four children, a boy and a girl, develop cataracts. What is the chance that their next child will be a girl with cataracts?
4. In cats, black coat (*B*) is codominant with yellow (*b*). The coat color gene is on the X chromosome. Calico cats, which have coats with black and yellow patches, are heterozygous for the coat color alleles.

- a. Why are most calico cats females?
- b. A calico female, Pickle, had a litter with one yellow male, two black males, two yellow females and three calico females. What were the genotype and phenotype of the father?

5. In *Drosophila*, three autosomal genes have alleles as follows: Gray body color (*G*) is dominant over black (*g*) Full wings (*A*) is dominant over vestigial (*a*) Red eye (*R*) is dominant over sepia (*r*)

Two crosses were performed, with the following results:

Cross I: Parents: heterozygous red, full × sepia, vestigial
Offspring: 131 red, full
120 sepia, vestigial
122 red, vestigial
127 sepia, full

Cross II: Parents: heterozygous gray, full × black, vestigial
Offspring: 236 gray, full
253 black, vestigial
50 gray, vestigial
61 black, full

Are any of the three genes linked on the same chromosome? If so, what is the map distance between the linked genes?

6. In a particular plant species, two alleles control flower color, which can be yellow, blue, or white. Crosses of these plants produce the following offspring:

Parental phenotypes	Offspring phenotypes (ratio)
Yellow × yellow	All yellow
Blue × yellow	Blue or yellow (1:1)
Blue × white	Blue or white (1:1)
White × white	All white

What will be the phenotype, and ratio, of the offspring of a cross of blue × blue?

7. In *Drosophila melanogaster*, the recessive allele *p*, when homozygous, determines pink eyes. *Pp* or *PP* results in wild-type eye color. Another gene on a different chromosome has a recessive allele, *sw*, that produces short wings when homozygous. Consider a cross between females of genotype *PPSwSw* and males of genotype *ppswsw*. Describe the phenotypes and genotypes of the *F*₁ generation and of the *F*₂ generation, produced by allowing the *F*₁ progeny to mate with one another.

8. On the same chromosome of *Drosophila melanogaster* that carries the *p* (pink eyes) locus, there is another locus that affects the wings. Homozygous recessives, *byby*, have blistery wings, while the dominant allele *By* produces wild-type wings. The *P* and *By* loci are very close together on the chromosome; that is, the two loci are tightly linked. In answering Questions 8a and 8b, assume that no crossing over occurs, and that the *F*₂ generation is produced by interbreeding the *F*₁ progeny.

- a. For the cross *PPByBy* × *ppbyby*, give the phenotypes and genotypes of the *F*₁ and *F*₂ generations.
b. For the cross *PPbyby* × *ppByBy*, give the phenotypes and genotypes of the *F*₁ and *F*₂ generations.
c. For the cross of Question 8b, what further phenotype(s) would appear in the *F*₂ generation if crossing over occurred?

- d. Draw a nucleus undergoing meiosis at the stage in which the crossing over (Question 8c) occurred. In which generation (*P*, *F*₁, or *F*₂) did this crossing over take place?

9. In chickens, when the dominant alleles of the genes for rose comb (*R*) and pea comb (*A*) are present together (*R_A_*), the result is a bird with a walnut comb. Chickens that are homozygous recessive for both genes produce a single comb. A rose-combed bird mated with a walnut-combed bird and produced offspring in the proportion:
3/8 walnut:3/8 rose:1/8 pea:1/8 single
What were the genotypes of the parents?

10. In *Drosophila melanogaster*, white (*w*), eosin (*w^e*), and wild-type red (*w⁺*) are multiple alleles at a single locus for eye color. This locus is on the X chromosome. A female that has eosin (pale orange) eyes is crossed with a male that has wild-type eyes. All the female progeny are red-eyed; half the male progeny have eosin eyes, and half have white eyes.
a. What is the order of dominance of these alleles?
b. What are the genotypes of the parents and progeny?

11. In humans, red-green color blindness is determined by an X-linked recessive allele (*a*), while eye color is determined by an autosomal gene, where brown (*B*) is dominant over blue (*b*).
 - a. What gametes can be formed with respect to these genes by a heterozygous, brown-eyed, color-blind male?
 - b. If a blue-eyed mother with normal vision has a brown-eyed, color-blind son and a blue-eyed, color-blind daughter, what are the genotypes of both parents and children?
12. If the dominant allele *A* is necessary for hearing in humans, and another allele, *B*, located on a different chromosome, results in deafness no matter what other genes are present, what percentage of the offspring of the marriage of *aaBb* × *Aabb* will be deaf?
13. The disease Leber's optic neuropathy is caused by a mutation in a gene carried on mitochondrial DNA. What would be the phenotype of their first child if a man with this disease married a woman who did not have the disease? What would be the result if the wife had the disease and the husband did not?

ADDITIONAL INVESTIGATION

Sometimes scientists get lucky. Consider Mendel's dihybrid cross shown in Figure 12.7. Peas have a haploid number of seven chromosomes, so many of their genes are linked. What would Mendel's results have been if the genes for seed color

and seed shape were linked with a map distance of ten units? Now, consider Morgan's fruit flies (see Figure 12.21). Suppose that the genes for body color and wing shape were not linked? What results would Morgan have obtained?

WORKING WITH DATA (GO TO yourBioPortal.com)

Mendel's Monohybrid Experiments Mendel's experiments with pea plants (Figure 12.3) laid the foundations of genetics. In this real-world exercise, you will analyze Mendel's data

from his published paper and see how he came to his conclusions about the nature of genes.

13

DNA and Its Role in Heredity

A structure for our times

Jurassic Park, in both its literary and film incarnations, features a fictional theme park populated with live dinosaurs. In the story, scientists isolate DNA from dinosaur blood extracted from the digestive tracts of fossil insects. The insects supposedly sucked the reptiles' blood right before being preserved in amber (fossilized tree resin). This DNA, according to the novel, could be manipulated to produce living individuals of long-extinct organisms such as velociraptors and the ever-memorable *Tyrannosaurus rex*.

The late Michael Crichton got the idea for his novel from an actual scientific paper in which the authors claimed to have detected reptilian DNA sequences in a fossil insect. Unfortunately, upon additional study, the "preserved" DNA turned out to be a contaminant from modern organisms.



Despite the facts that (1) the preservation of intact DNA over millions of years is highly improbable, and (2) DNA alone cannot generate a new organism, the huge success of Crichton's book brought DNA to the attention of millions. But even before *Jurassic Park*, the DNA double helix was a familiar secular icon.

The double helix first appeared in 1953, in a short paper by James Watson and Francis Crick in the journal *Nature*. An illustration of the molecule's structure drawn by Crick's wife, Odile, accompanied the article, and its simplicity and elegance caught the imagination of the general public as well as the intellect of scientists. As Watson later put it, "A structure this pretty just had to exist."

The double-helical structure of *deoxyribonucleic acid* is perhaps the most widely recognized symbol of modern science, and "DNA" has become part of everyday speech. One sees advertisements for a company whose customers get "into the DNA of business."

A digital media software system is called the "DNA Server." A perfume called DNA bills itself as "the essence of life."

Salvador Dali was the first well-known artist to use the DNA double helix in his whimsical creations in 1958. Today, sculptures representing the DNA double helix abound, and it is not only DNA's appearance that stirs our imagination. The DNA nucleotide sequence itself, the "code for life," has inspired unique works of art that incorporate real DNA molecules. A portrait of Sir John Sulston, a Nobel prize-winning geneticist, is made of tiny bacterial colonies, each containing a piece of Sulston's DNA. The Brazilian artist Eduardo Kac translated a sentence from the Bible into Morse code, and from Morse

Reviving the Velociraptor Scientists and artists have been creating inanimate reconstructions of dinosaurs for more than 100 years. Michael Crichton's novel *Jurassic Park* was based on the fictional premise that DNA retrieved from fossils could produce living dinosaurs, such as this velociraptor.



In the Nature of Things The double helix of DNA has become an iconic symbol of modern science and culture. Artists and designers make use of the widely recognized shape in many ways.

code into a DNA sequence. The sequence was synthesized and incorporated into bacteria. Viewers could turn on an ultraviolet lamp to create mutations in the DNA (and thus in the biblical verse it encoded).

For many people, DNA has come to symbolize the promise and perils of our rapidly expanding knowledge of genetics. Although DNA sequences alone cannot generate a new organism, *biotechnologies* using DNA can modify existing organisms into essentially new organisms. As we will see in Chapter 18, such use of this iconic molecule has generated both excitement and concern about potential risks.

IN THIS CHAPTER we will describe the key experiments that led to the identification of DNA as the genetic material. We will then describe the structure of the DNA molecule and how this structure determines its function. We will describe the processes by which DNA is replicated, repaired, and maintained. Finally, we present an important practical application arising from our knowledge of DNA replication: the polymerase chain reaction.

CHAPTER OUTLINE

- 13.1 What Is the Evidence that the Gene Is DNA?
- 13.2 What Is the Structure of DNA?
- 13.3 How Is DNA Replicated?
- 13.4 How Are Errors in DNA Repaired?
- 13.5 How Does the Polymerase Chain Reaction Amplify DNA?

13.1 What Is the Evidence that the Gene is DNA?

By the early twentieth century, geneticists had associated the presence of genes with chromosomes. Research began to focus on exactly which chemical component of chromosomes comprised this genetic material.

By the 1920s, scientists knew that chromosomes were made up of DNA and proteins. At this time a new dye was developed by Robert Feulgen that could bind specifically to DNA and that stained cell nuclei red in direct proportion to the amount of DNA present in the cell. This technique provided circumstantial evidence that DNA was the genetic material:

- *It was in the right place.* DNA was confirmed to be an important component of the nucleus and the chromosomes, which were known to carry genes.
- *It varied among species.* When cells from different species were stained with the dye and their color intensity measured, each species appeared to have its own specific amount of nuclear DNA.
- *It was present in the right amounts.* The amount of DNA in somatic cells (body cells not specialized for reproduction) was twice that in reproductive cells (eggs or sperm)—as might be expected for diploid and haploid cells, respectively.

But circumstantial evidence is *not* a scientific demonstration of cause and effect. After all, proteins are also present in cell nuclei. Science relies on experiments to test hypotheses. The convincing demonstration that DNA is the genetic material came from two sets of experiments, one on bacteria and the other on viruses.

DNA from one type of bacterium genetically transforms another type

The history of biology is filled with incidents in which research on one specific topic has—with or without answering the question originally asked—contributed richly to another, apparently unrelated area. Such a case of serendipity is seen in the work of Frederick Griffith, an English physician.

In the 1920s, Griffith was studying the bacterium *Streptococcus pneumoniae*, or pneumococcus, one of the agents that cause pneumonia in humans. He was trying to develop a vaccine against this devastating illness (antibiotics had not yet been discovered). Griffith was working with two strains of pneumococcus:

- Cells of the S strain produced colonies that looked smooth (S). Covered by a polysaccharide capsule, these cells were

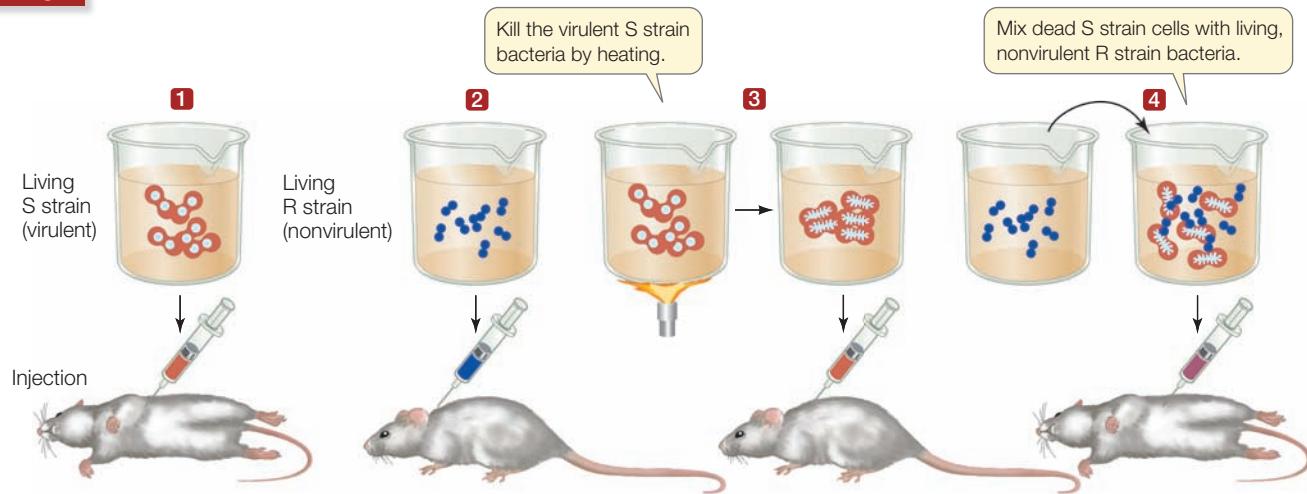
INVESTIGATING LIFE

13.1 Genetic Transformation

Griffith's experiments demonstrated that something in the virulent S strain of pneumococcus could transform nonvirulent R strain bacteria into a lethal form, even when the S strain bacteria had been killed by high temperatures.

HYPOTHESIS Material in dead bacterial cells can genetically transform living bacterial cells.

METHOD



RESULTS

1 Mouse dies

Living S strain cells found in heart

2 Mouse healthy

No bacterial cells found in heart

3 Mouse healthy

No bacterial cells found in heart

4 Mouse dies

Living S strain cells found in heart

CONCLUSION A chemical substance from one cell is capable of genetically transforming another cell.

FURTHER INVESTIGATION: How would you show that heat-killed R strain bacteria can transform living S strain bacteria?

Go to yourBioPortal.com for original citations, discussions, and relevant links for all INVESTIGATING LIFE figures.

protected from attack by a host's immune system. When S cells were injected into mice, they reproduced and caused pneumonia (the strain was *virulent*).

- Cells of the R strain produced colonies that looked rough (R), lacked the protective capsule, and were not virulent.

Griffith inoculated some mice with heat-killed S-type pneumococcus cells. These heat-killed bacteria did not produce infection. However, when Griffith inoculated other mice with a mixture of living R bacteria and heat-killed S bacteria, to his astonishment, the mice died of pneumonia (Figure 13.1). When he examined blood from the hearts of these mice, he found it full of living bacteria—many of them with characteristics of the virulent S strain! Griffith concluded that in the presence of the dead S-type pneumococcus cells, some of the living R-type cells had been transformed into virulent S cells. The fact that these

S-type cells reproduced to make more S-type cells showed that the change from R-type to S-type was genetic.

Did this transformation of the bacteria depend on something that happened in the mouse's body? No. It was shown that simply incubating living R and heat-killed S bacteria together in a test tube yielded the same transformation. Years later, another group of scientists discovered that a cell-free extract of heat-killed S cells could also transform R cells. (A cell-free extract contains all the contents of ruptured cells, but no intact cells.) This result demonstrated that some substance—called at the time a chemical **transforming principle**—from the dead S pneumococcus cells could cause a heritable change in the affected R cells. This was an extraordinary discovery: treatment with a chemical substance could permanently change an inherited characteristic. Now it remained to identify the chemical structure of this substance.

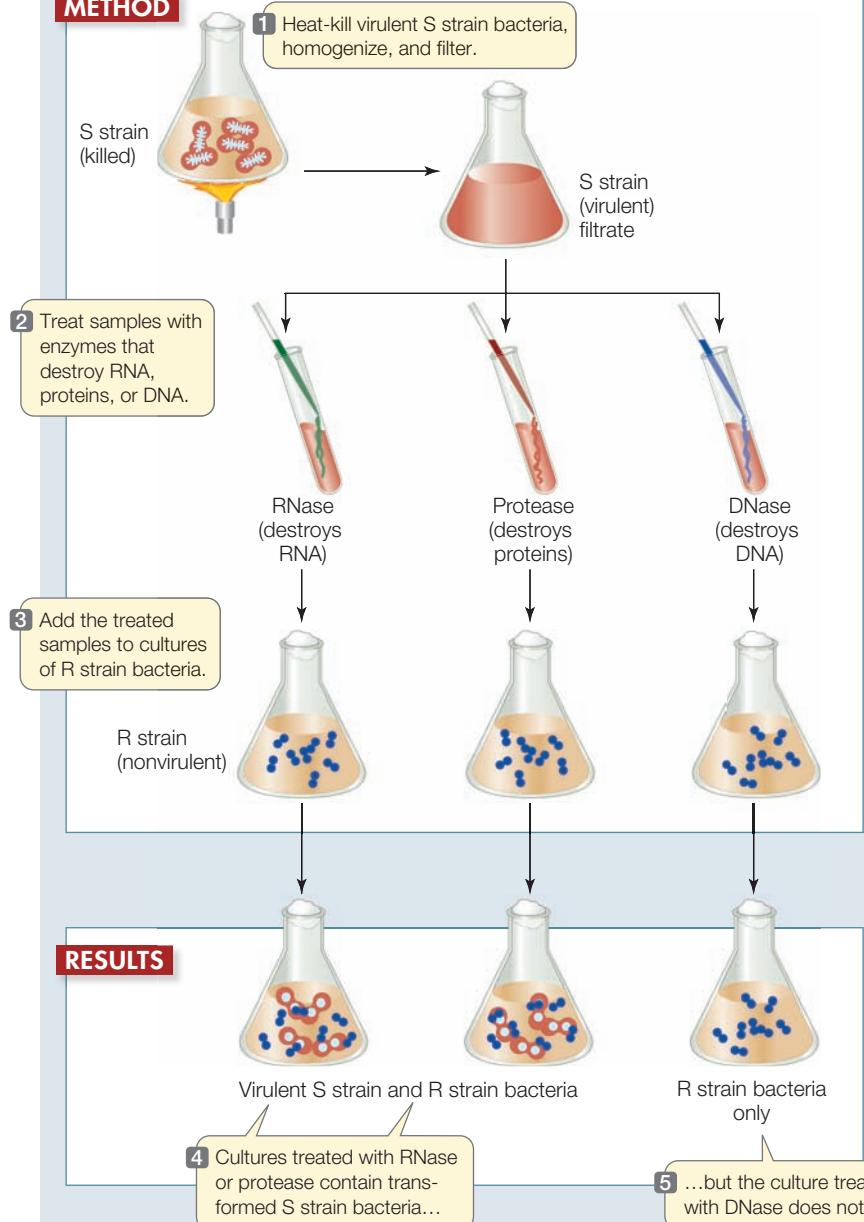
INVESTIGATING LIFE

13.2 Genetic Transformation by DNA

Experiments by Avery, MacLeod, and McCarty showed that DNA from the virulent S strain of pneumococcus was responsible for the transformation in Griffith's experiments (see Figure 13.1).

HYPOTHESIS The chemical nature of the transforming substance from pneumococcus is DNA.

METHOD



CONCLUSION Because only DNase destroyed the transforming substance, the transforming substance is DNA.

Go to yourBioPortal.com for original citations, discussions, and relevant links for all INVESTIGATING LIFE figures.

The transforming principle is DNA

Identifying the transforming principle was a crucial step in the history of biology. Work on identifying the transforming principle was completed by Oswald Avery and his colleagues at what is now The Rockefeller University.

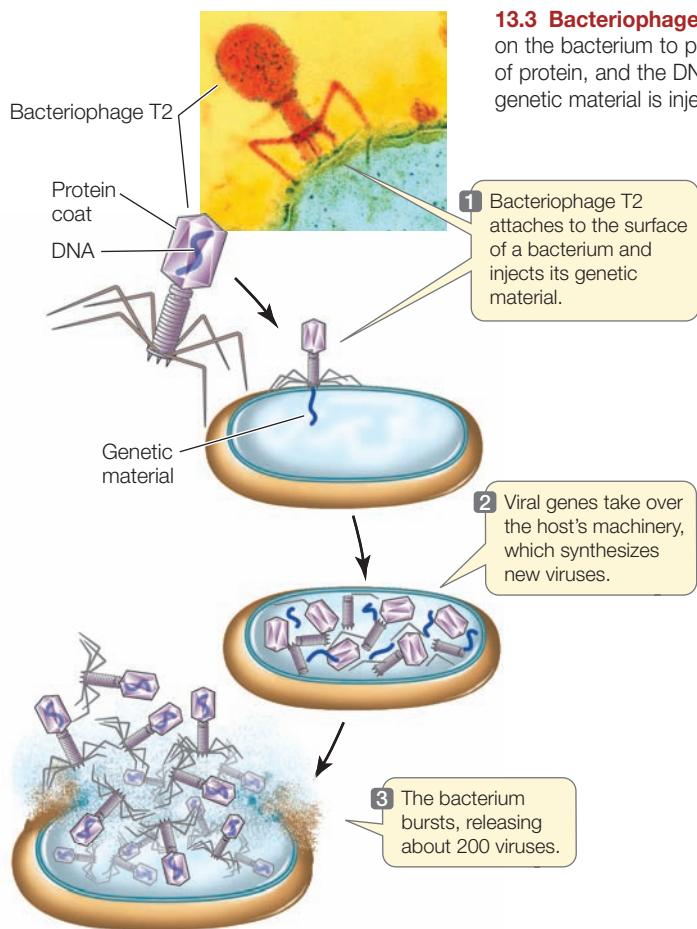
They treated samples known to contain the pneumococcal transforming principle in a variety of ways to destroy different types of molecules—proteins, nucleic acids, carbohydrates, and lipids—and tested the treated samples to see if they had retained the transforming activity. The answer was always the same: if the DNA in the sample was destroyed, transforming activity was lost, but there was no loss of activity when proteins, carbohydrates, or lipids were destroyed (Figure 13.2). As a final step, Avery and his colleagues Colin MacLeod and Maclyn McCarty isolated virtually pure DNA from a sample containing pneumococcal-transforming principle, and showed that it caused bacterial transformation. We now know that the gene for the enzyme that catalyzes the synthesis of the polysaccharide capsule, which makes the bacteria look “smooth,” was transferred during transformation.

Genetic transformation occurs in nature, although only in certain species of bacteria such as *Pneumococcus*. It does not occur, for example, in *E. coli*. Cells can pick up DNA fragments released into the environment by dead and ruptured cells. Only a small part of the genome is taken up by the transformed cells. Once the new DNA enters the cell, a transforming event very similar to recombination occurs (see Figure 12.26), and new genes can be incorporated into the host chromosome.

The work of Avery's group was a milestone in establishing that DNA is the genetic material in bacterial cells. However, when it was first published in 1944, it had little impact, for two reasons. First, most scientists did not believe that DNA was chemically complex enough to be the genetic material, especially given the much greater chemical complexity of proteins. Second, and perhaps more important, bacterial genetics was a new field of study—it was not yet clear that bacteria even had genes.

Viral replication experiments confirmed that DNA is the genetic material

The questions about bacteria and other simple organisms were soon resolved, as researchers identified genes and mutations. Bacteria and viruses seemed to undergo genetic processes



13.3 Bacteriophage T2: Reproduction Cycle Bacteriophage T2 is parasitic on *E. coli*, depending on the bacterium to produce new viruses. The external structures of bacteriophage T2 consist entirely of protein, and the DNA is contained within the protein coat. When the virus infects an *E. coli* cell, its genetic material is injected into the host bacterium.

- Proteins contain some sulfur (in the amino acids cysteine and methionine). Sulfur is not present in DNA, and has a radioactive isotope, ^{35}S . Hershey and Chase grew bacteriophage T2 in a bacterial culture in the presence of ^{35}S , so the proteins of the resulting viruses were labeled with (contained) the radioisotope.
- DNA contains phosphorus (in the deoxyribose-phosphate backbone—see Figure 4.2). Phosphorus is not present in most proteins, and it also has a radioisotope, ^{32}P . The researchers grew another batch of T2 in a bacterial culture in the presence of ^{32}P , thus labeling the viral DNA with ^{32}P .

Using these radioactively labeled viruses, Hershey and Chase performed their revealing experiments (Figure 13.4). In one experiment, they allowed ^{32}P -labeled bacteriophage to infect bacteria; in the other, the bacteria were infected by ^{35}S -labeled bacteriophage. After a few minutes, they agitated each mixture of infected bacteria vigorously in a kitchen blender, which stripped away the parts of the virus that had not penetrated the bacteria, without bursting the bacteria. Then they separated the bacteria from the rest of the material in a centrifuge.

Spinning solutions or suspensions at high speed in a centrifuge causes the solutes and/or particles to separate and form a gradient according to their densities. The lighter remains of the viruses (those parts that had not penetrated the bacteria) were captured in the “supernatant” fluid, while the heavier bacterial cells segregated into a “pellet” in the bottom of the centrifuge tube. The scientists found that the supernatant fluid contained most of the ^{35}S (and thus the viral protein), while most of the ^{32}P (and thus the viral DNA) had stayed with the bacteria. These results suggested that it was DNA that had been transferred into the bacteria, and that DNA was the compound responsible for redirecting the genetic program of the bacterial cell.

Hershey and Chase performed similar but longer-term experiments, allowing the progeny (offspring) generation of viruses to grow. The resulting viruses contained almost no ^{35}S and none of the parental viral protein. They did, however, contain about one-third of the original ^{32}P —and thus, presumably, one-third of the original DNA. Because DNA was carried over in the viruses from generation to generation but protein was not, the logical conclusion was that the hereditary information was contained in the DNA.

Eukaryotic cells can also be genetically transformed by DNA

With the publication of the evidence for DNA as the genetic material in bacteria and viruses, the question arose as to whether DNA was also the genetic material in complex eukaryotes. Some dubious experimental results were reported. For example, a

similar to those in fruit flies and pea plants. Experiments with these relatively simple organisms were designed to discover the nature of the genetic material.

In 1952, Alfred Hershey and Martha Chase of the Carnegie Laboratory of Genetics published a paper that had a much greater immediate impact than Avery's 1944 paper. The Hershey–Chase experiment, which sought to determine whether DNA or protein was the genetic material, was carried out with a virus that infects bacteria. This virus, called bacteriophage T2, consists of little more than a DNA core packed inside a protein coat (Figure 13.3). Thus the virus is made of the two materials that were, at the time, the leading candidates for the genetic material.

When bacteriophage T2 attacks a bacterium, part (but not all) of the virus enters the bacterial cell. About 20 minutes later, the cell bursts, releasing dozens of particles that are virtually identical to the infecting virus particle. Clearly the virus is somehow able to replicate itself inside the bacterium. Hershey and Chase deduced that the entry of some viral component affects the genetic program of the host bacterial cell, transforming it into a bacteriophage factory. They set out to determine which part of the virus—protein or DNA—enters the bacterial cell. To trace the two components of the virus over its life cycle, Hershey and Chase labeled each component with a specific radioisotope:

INVESTIGATING LIFE

13.4 The Hershey–Chase Experiment

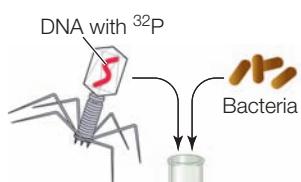
When bacterial cells were infected with radioactively labeled T2 bacteriophage, only labeled DNA was found in the bacteria. After centrifuging the culture to make the bacteria form a pellet, the labeled protein remained in the supernatant. This showed that DNA, not protein, is the genetic material.

HYPOTHESIS Either component of a bacteriophage—DNA or protein—might be the hereditary material that enters a bacterial cell to direct the assembly of new viruses.

METHOD

Experiment 1

1a Label phage. P is an element in DNA, but not in proteins.



2 Infect bacteria with labeled viruses.



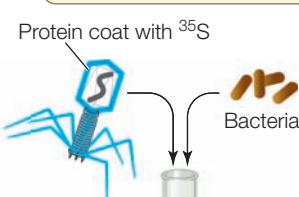
3 Agitate in a blender to detach viruses from bacterial cells.



4 Centrifuge to force the bacterial cells to the bottom of the tube, forming a pellet. Supernatant fluid contains the viruses.

Experiment 2

1b Label phage. S is an element in proteins, but not in DNA.



3 Agitate in a blender to detach viruses from bacterial cells.



4 Centrifuge to force the bacterial cells to the bottom of the tube, forming a pellet. Supernatant fluid contains the viruses.

RESULTS

5a Most of the ^{32}P is in the pellet with the bacteria.



5b Most of the ^{35}S is in the supernatant fluid with the viruses.



CONCLUSION DNA, not protein, enters bacterial cells and directs the assembly of new viruses.

Go to yourBioPortal.com for original citations, discussions, and relevant links for all INVESTIGATING LIFE figures.

white duck was injected with DNA from a brown duck and the recipient was reported to turn brown. In another example, flatworms were fed DNA from worms that had learned a simple task, and the recipient worms were reported to immediately get smarter. However, no one could duplicate these results. This episode underscores a central aspect of experimental biology: that published research should be repeated with the same results before the conclusions can be considered valid.

It would be impossible for a large molecule such as DNA to avoid hydrolysis into nucleotides in the digestive system, let alone get into all the cells of the body, after being ingested by an animal. However, genetic transformation of eukaryotic cells by DNA (called **transfection**) can be demonstrated. The key is to use a **genetic marker**, a gene whose presence in the recipient cell confers an observable phenotype. In the experiments with pneumococcus, these phenotypes were the smooth polysaccharide capsule and virulence. In eukaryotes, researchers usually use a nutritional or antibiotic resistance marker gene that permits the growth of transformed recipient cells but not of nontransformed cells. For example, thymidine kinase is an enzyme needed to make use of thymidine in the synthesis of deoxythymidine triphosphate (dTTP), one of the four deoxyribonucleoside triphosphates used in the synthesis of DNA. Mammalian cells that lack the gene for thymidine kinase cannot grow in a medium that contains thymidine as the only source for dTTP synthesis. When DNA containing the marker gene encoding thymidine kinase is added to a culture of mammalian cells lacking this gene, some cells will grow in the thymidine medium, demonstrating that they have been transfected with the gene (Figure 13.5). Any cell can be transfected in this way, even an egg cell. In this case, a whole new genetically transformed organism can result; such an organism is referred to as *transgenic*. Transformation in eukaryotes is the final line of evidence for DNA as the genetic material.

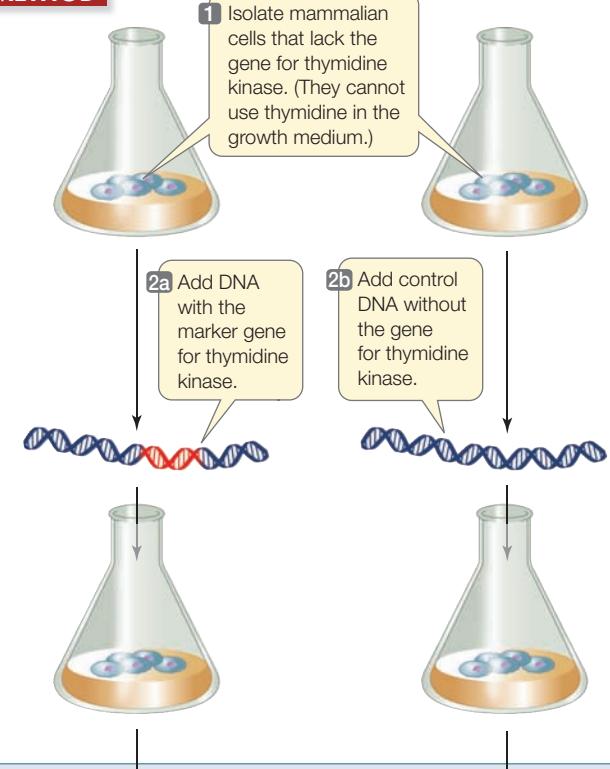
INVESTIGATING LIFE

13.5 Transfection in Eukaryotic Cells

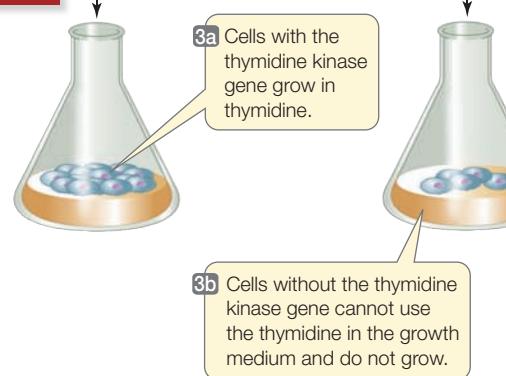
The use of a marker gene shows that mammalian cells can be genetically transformed by DNA. Usually, the marker gene is carried by a larger molecule (a virus or a small chromosome).

HYPOTHESIS DNA can transform eukaryotic cells.

METHOD



RESULTS



CONCLUSION The cells were transformed by DNA.

Go to yourBioPortal.com for original citations, discussions, and relevant links for all INVESTIGATING LIFE figures.

13.1 RECAP

Experiments on bacteria and on viruses demonstrated that DNA is the genetic material.

- At the time of Griffith's experiments in the 1920s, what circumstantial evidence suggested to scientists that DNA might be the genetic material? See p. 267
- Why were the experiments of Avery, MacLeod, and McCarty definitive evidence that DNA was the genetic material? See p. 269 and Figure 13.2
- What attributes of bacteriophage T2 were key to the Hershey–Chase experiments demonstrating that DNA is the genetic material? See p. 270 and Figure 13.4

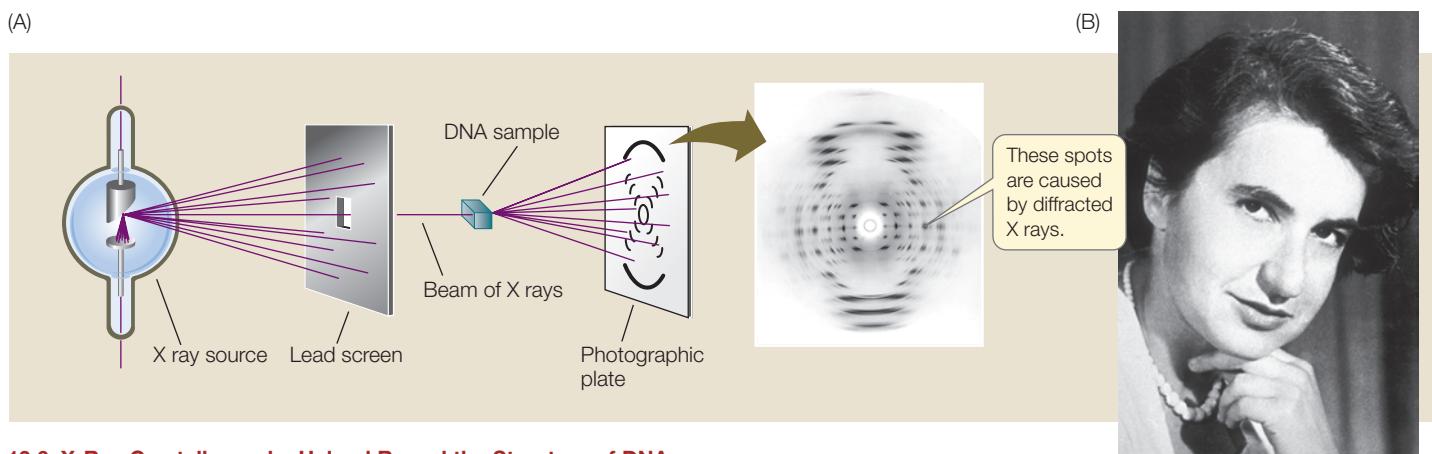
As soon as scientists became convinced that the genetic material was DNA, they began efforts to learn its precise three-dimensional chemical structure. The chemical makeup of DNA, as a polymer made up of nucleotide monomers, had been known for several decades. In determining the structure of DNA, scientists hoped to find the answers to two questions: (1) how is DNA replicated between cell divisions, and (2) how does it direct the synthesis of specific proteins? They were eventually able to answer both questions.

13.2 What Is the Structure of DNA?

The structure of DNA was deciphered only after many types of experimental evidence were considered together in a theoretical framework. The most crucial evidence was obtained using X-ray crystallography. Some chemical substances, when they are isolated and purified, can be made to form crystals. The positions of atoms in a crystallized substance can be inferred from the diffraction pattern of X rays passing through the substance (Figure 13.6A). The structure of DNA would not have been characterized without the crystallographs prepared in the early 1950s by the English chemist Rosalind Franklin (Figure 13.6B). Franklin's work, in turn, depended on the success of the English biophysicist Maurice Wilkins, who prepared samples containing very uniformly oriented DNA fibers. These DNA samples were far better for diffraction than previous ones, and the crystallographs Franklin prepared from them suggested a spiral or helical molecule.

The chemical composition of DNA was known

The chemical composition of DNA also provided important clues to its structure. Biochemists knew that DNA was a polymer of nucleotides. Each nucleotide consists of a molecule of the sugar deoxyribose, a phosphate group, and a nitrogen-containing base (see Figures 4.1 and 4.2). The only differences among the four nucleotides of DNA are their nitrogenous bases: the purines **adenine** (A) and **guanine** (G), and the pyrimidines **cytosine** (C) and **thymine** (T).



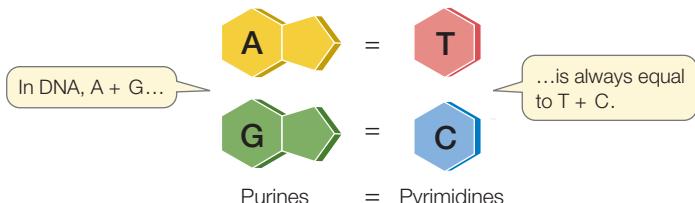
13.6 X-Ray Crystallography Helped Reveal the Structure of DNA

(A) The positions of atoms in a crystallized chemical substance can be inferred by the pattern of diffraction of X rays passed through it. The pattern of DNA is both highly regular and repetitive. (B) Rosalind Franklin's crystallography helped scientists to visualize the helical structure of the DNA molecule.

In 1950, biochemist Erwin Chargaff at Columbia University reported some observations of major importance. He and his colleagues found that DNA from many different species—and from different sources within a single organism—exhibits certain regularities. In almost all DNA, the following rule holds: The amount of adenine equals the amount of thymine ($A = T$), and the amount of guanine equals the amount of cytosine ($G = C$) (Figure 13.7). As a result, the total abundance of purines ($A + G$) equals the total abundance of pyrimidines ($T + C$). The structure of DNA could not have been worked out without this observation, now known as Chargaff's rule, yet its significance was overlooked for at least three years.

Watson and Crick described the double helix

The solution to the structure of DNA was finally achieved through model building: the assembly of three-dimensional representations of possible molecular structures using known relative molecular dimensions and known bond angles. This technique was originally applied to molecular structural studies by the American biochemist Linus Pauling. The English physicist Francis Crick and the American geneticist James D. Watson (Figure 13.8A), who were both then at the Cavendish Laboratory of Cambridge University, used model building to solve the structure of DNA.

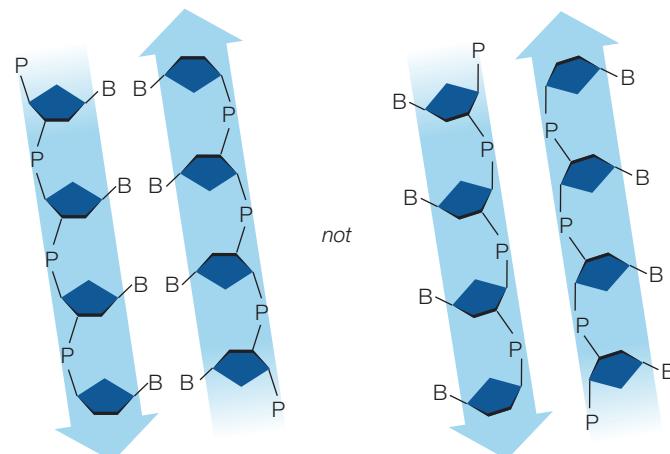


13.7 Chargaff's Rule In DNA, the total abundance of purines is equal to the total abundance of pyrimidines.

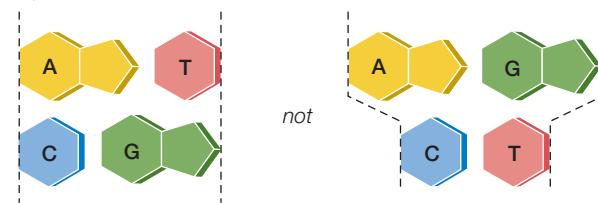
Watson and Crick attempted to combine all that had been learned so far about DNA structure into a single coherent model. Rosalind Franklin's crystallography results (see Figure 13.6) convinced Watson and Crick that the DNA molecule must be **helical** (cylindrically spiral). Density measurements and previous model building results suggested that there are two polynucleotide chains in the molecule. Modeling studies also showed that the strands run in opposite directions, that is, they are **antiparallel**; that two strands would not fit together in the model if they were parallel.

How are the nucleotides oriented in these chains? Watson and Crick suggested that:

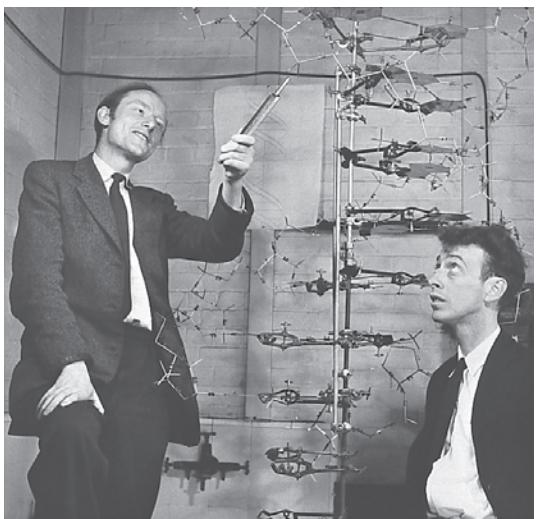
- The nucleotide bases are on the interior of the two strands, with a sugar-phosphate backbone on the outside:



- To satisfy Chargaff's rule (purines = pyrimidines), a purine on one strand is always paired with a pyrimidine on the opposite strand. These **base pairs** ($A-T$ and $G-C$) have the same width down the double helix, a uniformity shown by x-ray diffraction.



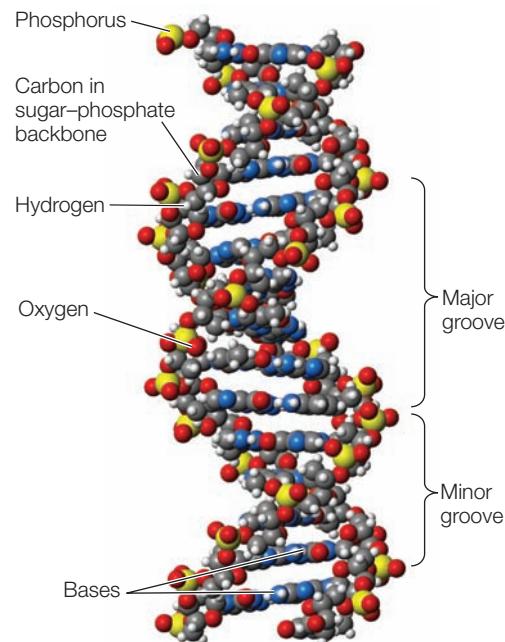
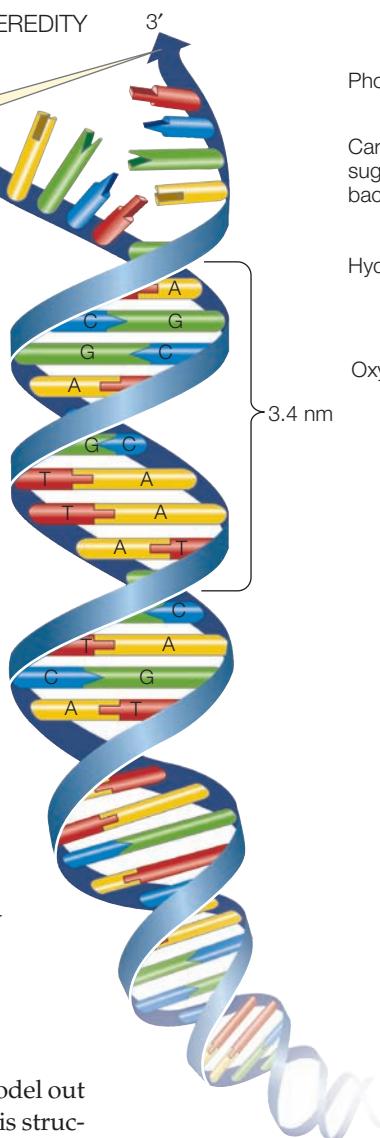
(A)



(B)

The blue bands represent the two sugar-phosphate backbones, which run in opposite directions:

5' 3'
3' 5'



13.8 DNA Is a Double Helix (A) Francis Crick (left) and James Watson (right) proposed that the DNA molecule has a double-helical structure. (B) Biochemists can now pinpoint the position of every atom in a DNA molecule. To see that the essential features of the original Watson-Crick model have been verified, follow with your eyes the double-helical chains of sugar-phosphate groups and note the horizontal rungs of the bases.

In late February of 1953, Crick and Watson built a model out of tin that established the general structure of DNA. This structure explained all the known chemical properties of DNA, and it opened the door to understanding its biological functions. There have been minor amendments to that first published structure, but its principal features remain unchanged.

Four key features define DNA structure

Four features summarize the molecular architecture of the DNA molecule (see **Figure 13.8B**):

- It is a *double-stranded helix* of uniform diameter.
- It is *right-handed*. (Hold your right hand with the thumb pointing up. Imagine the curve of the helix following the direction of your fingers as it winds upward and you have the idea.)
- It is *antiparallel* (the two strands run in opposite directions).
- The outer edges of the nitrogenous bases are *exposed* in the major and minor grooves. These grooves exist because the backbones of the two strands are closer together on one side of the double helix (forming the minor groove) than on the other side (forming the major groove).

THE HELIX The sugar-phosphate “backbones” of the polynucleotide chains coil around the outside of the helix, and the ni-

trogenous bases point toward the center. The two chains are held together by hydrogen bonding between specifically paired bases (**Figure 13.9**). Consistent with Chargaff’s rule,

- Adenine (A) pairs with thymine (T) by forming two hydrogen bonds.
- Guanine (G) pairs with cytosine (C) by forming three hydrogen bonds.

Every base pair consists of one purine (A or G) and one pyrimidine (T or C). This pattern is known as **complementary base pairing**.

Because the A-T and G-C pairs are of equal length, they fit into a fixed distance between the two chains (like rungs on a ladder), and the diameter of the helix is thus uniform. The base pairs are flat, and their stacking in the center of the molecule is stabilized by hydrophobic interactions (see Section 2.2), contributing to the overall stability of the double helix.

ANTIPARALLEL STRANDS What does it mean to say that the two DNA strands are *antiparallel*? The direction of each strand is determined by examining the bonds between the alternating phosphate and sugar groups that make up the backbone of each strand. Look closely at the five-carbon sugar (deoxyribose)

13.9 Base Pairing in DNA Is Complementary The purines (A and G) pair with the pyrimidines (T and C, respectively) to form base pairs that are equal in size and resemble the rungs on a ladder whose sides are formed by the sugar-phosphate backbones. The deoxyribose sugar (left) is where the 3' and 5' carbons are located. The two strands are antiparallel.

molecule in Figure 13.9. The number followed by a prime (') designates the position of a carbon atom in the sugar. In the sugar-phosphate backbone of DNA, the phosphate groups are connected to the 3' carbon of one deoxyribose molecule and the 5' carbon of the next, linking successive sugars together.

Thus the two ends of a polynucleotide chain differ. At one end of a chain is a free (not connected to another nucleotide) 5' phosphate group ($-\text{OPO}_3^-$); this is called the **5' end**. At the other end is a free 3' hydroxyl group ($-\text{OH}$); this is called the **3' end**. In a DNA double helix, the 5' end of one strand is paired with the 3' end of the other strand, and vice versa. In other words, if you drew an arrow for each strand running from 5' to 3', the arrows would point in opposite directions.

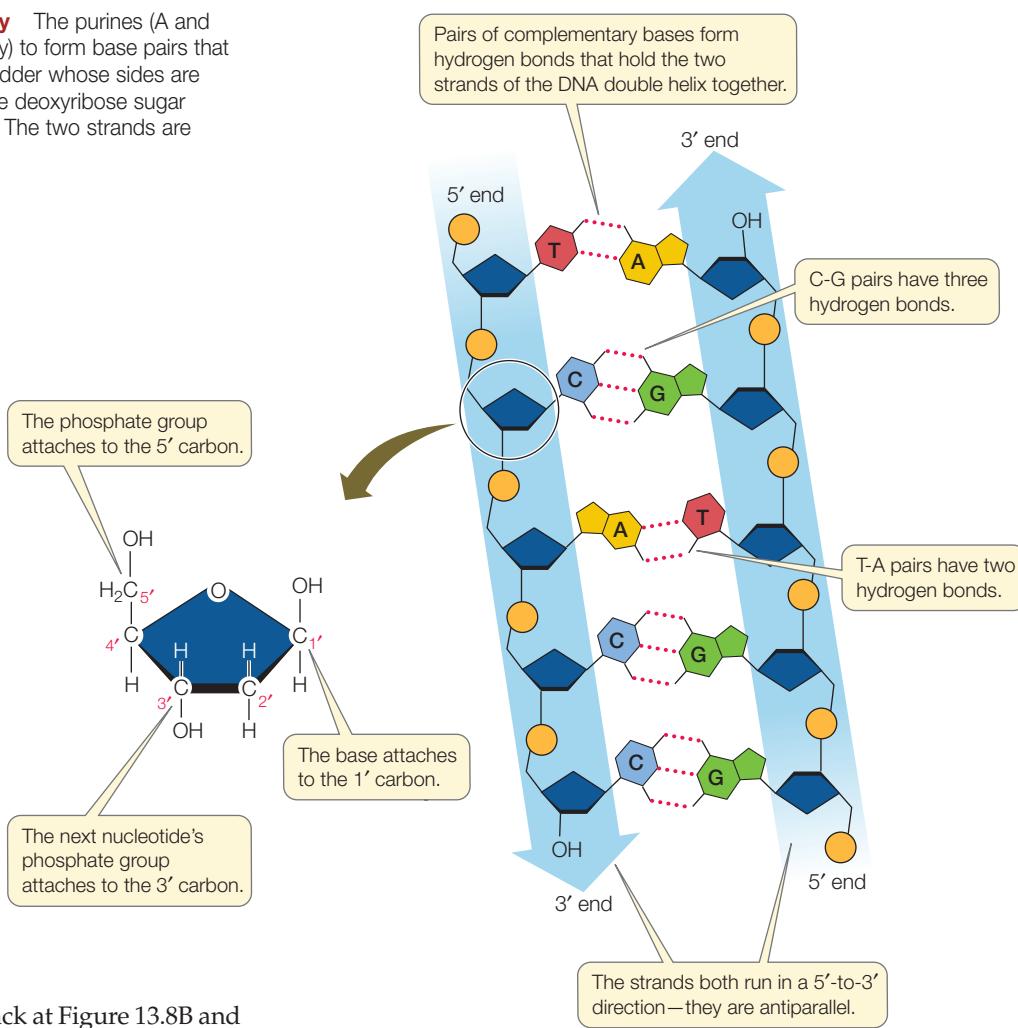
BASE EXPOSURE IN THE GROOVES Look back at Figure 13.8B and note the major and minor grooves in the helix. From these grooves, the exposed outer edges of the flat, hydrogen-bonded base pairs are accessible for additional hydrogen bonding. As seen in Figure 13.9, two hydrogen bonds join each A-T base pair, while three hydrogen bonds join each G-C base pair. Hydrogen-bonding opportunities also exist at an unpaired C=O group in T and an "N" in A. The G-C base pair offers additional hydrogen bonding possibilities as well. Thus the surfaces of the A-T and G-C base pairs are chemically distinct, allowing other molecules, such as proteins, to recognize specific base pair sequences and bind to them. Access to the exposed base-pair sequences in the major and minor grooves is the key to protein-DNA interactions, which are necessary for the replication and expression of the genetic information in DNA.

The double-helical structure of DNA is essential to its function

The genetic material performs four important functions, and the DNA structure proposed by Watson and Crick was elegantly suited to three of them.

- *The genetic material stores an organism's genetic information.*

With its millions of nucleotides, the base sequence of a DNA molecule can encode and store an enormous amount



of information. Variations in DNA sequences can account for species and individual differences. DNA fits this role nicely.

- *The genetic material is susceptible to mutations (permanent changes) in the information it encodes.* For DNA, mutations might be simple changes in the linear sequence of base pairs.
- *The genetic material is precisely replicated in the cell division cycle.* Replication could be accomplished by complementary base pairing, A with T and G with C. In the original publication of their findings in 1953, Watson and Crick coyly pointed out, "It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material."
- *The genetic material (the coded information in DNA) is expressed as the phenotype.* This function is not obvious in the structure of DNA. However, as we will see in the next chapter, the nucleotide sequence of DNA is copied into RNA, which uses the coded information to specify a linear sequence of amino acids—a protein. The folded forms of proteins determine many of the phenotypes of an organism.

13.2 RECAP

DNA is a double helix made up of two antiparallel polynucleotide chains. The two chains are joined by hydrogen bonds between the nucleotide bases, which pair specifically: A with T, and G with C. Chemical groups on the bases that are exposed in the grooves of the helix are available for hydrogen bonding with other molecules, such as proteins. These molecules can recognize specific sequences of nucleotide bases.

- Describe the evidence that Watson and Crick used to come up with the double helix model for DNA. **See p. 273**
- How does the double-helical structure of DNA relate to its function? **See p. 275**

Once the structure of DNA was understood, it was possible to investigate how DNA replicates itself. Let's examine the experiments that taught us how this elegant process works.

13.3 How Is DNA Replicated?

The mechanism of DNA replication that suggested itself to Watson and Crick was soon confirmed. First, experiments showed that DNA could be replicated in a test tube containing simple substrates and an enzyme. Then a truly classic experiment showed that each of the two strands of the double helix can serve as a template for a new strand of DNA.

yourBioPortal.com

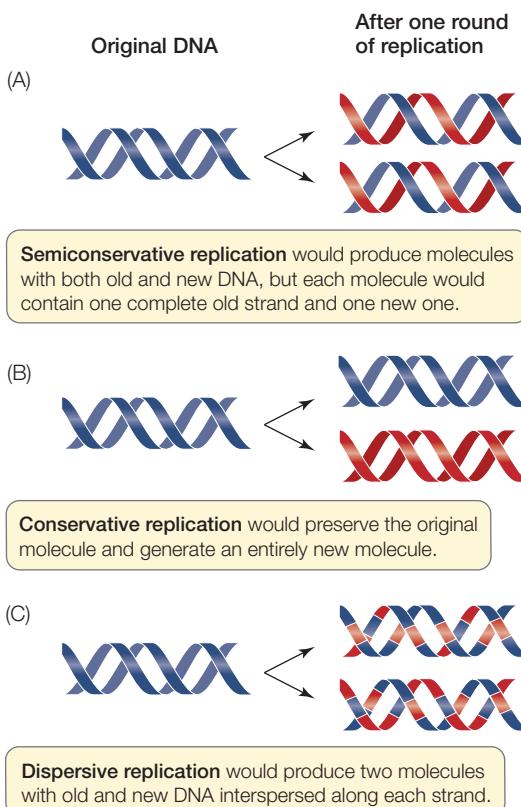
GO TO Animated Tutorial 13.1 • DNA Replication, Part 1: Replication of a Chromosome and DNA Polymerization

Three modes of DNA replication appeared possible

The prediction that the DNA molecule contains the information needed for its own replication was confirmed by the work of Arthur Kornberg, then at Washington University in St. Louis. He showed that new DNA molecules with the same base composition as the original molecules could be synthesized in a test tube containing the following substances:

- The substrates were the deoxyribonucleoside triphosphates dATP, dCTP, dGTP, and dTTP.
- A **DNA polymerase** enzyme catalyzed the reaction.
- DNA served as a **template** to guide the incoming nucleotides.
- The reaction also contained salts and a pH buffer, to create an appropriate chemical environment for the DNA polymerase to function.

Recall that a nucleoside is a nitrogen base attached to a sugar. The four deoxyribonucleoside triphosphates (dNTPs) each consist of a nitrogen base attached to deoxyribose, which in turn is attached to three phosphate groups. When a dNTP is added



13.10 Three Models for DNA Replication In each model, the original DNA is shown in blue and the newly synthesized DNA is in red.

to a DNA strand during DNA synthesis, the two terminal phosphates are removed, resulting in a monophosphate nucleotide.

The next challenge was to determine which of three possible replication patterns occurs during DNA replication:

- Semiconservative replication**, in which each parent strand serves as a template for a new strand, and the two new DNA molecules each have one old and one new strand (**Figure 13.10A**)
- Conservative replication**, in which the original double helix serves as a template for, but does not contribute to, a new double helix (**Figure 13.10B**)
- Dispersive replication**, in which fragments of the original DNA molecule serve as templates for assembling two new molecules, each containing old and new parts, perhaps at random (**Figure 13.10C**)

Watson and Crick's original paper suggested that DNA replication was semiconservative, but Kornberg's experiment did not provide a basis for choosing among these three models.

An elegant experiment demonstrated that DNA replication is semiconservative

The work of Matthew Meselson and Franklin Stahl convinced the scientific community that DNA is reproduced by **semiconservative replication**. Working at the California Institute of Technology, Meselson and Stahl devised a simple way to distinguish between old parent strands of DNA and newly copied ones: *density labeling*.

The key to their experiment was the use of a “heavy” isotope of nitrogen. Heavy nitrogen (^{15}N) is a rare, nonradioactive isotope that makes molecules containing it denser than chemically identical molecules containing the common isotope, ^{14}N . Meselson, Stahl, and Jerome Vinograd grew two cultures of the bacterium *Escherichia coli* for many generations:

- One culture was grown in a medium whose nitrogen source (ammonium chloride, NH_4Cl) was made with ^{15}N instead of ^{14}N . As a result, all the DNA in the bacteria was “heavy.”

yourBioPortal.com

GO TO Animated Tutorial 13.2 • The Meselson–Stahl Experiment

- Another culture was grown in a medium containing ^{14}N , and all the DNA in these bacteria was “light.”

When DNA extracts from the two cultures were combined and centrifuged, two separate bands formed, showing that this method could be used to distinguish between DNA samples of slightly different densities.

Next, the researchers grew another *E. coli* culture on ^{15}N medium, then transferred it to normal ^{14}N medium and allowed the bacteria to continue growing (Figure 13.11). Under the conditions they used, *E. coli* cells replicate their DNA and divide

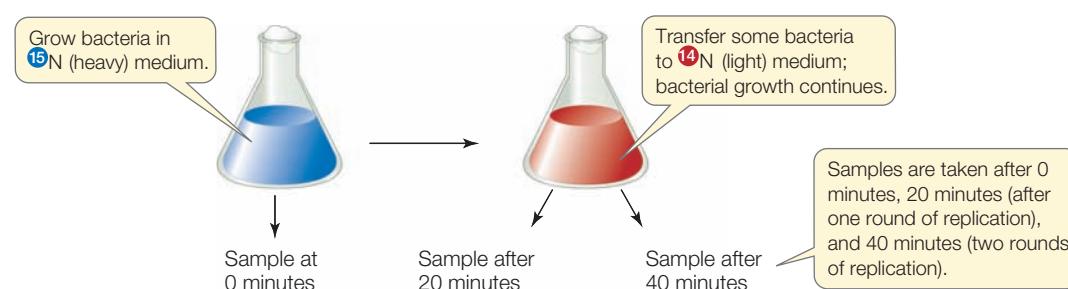
INVESTIGATING LIFE

13.11 The Meselson–Stahl Experiment

A centrifuge was used to separate DNA molecules labeled with isotopes of different densities. This experiment revealed a pattern that supports the semiconservative model of DNA replication.

HYPOTHESIS DNA replicates semiconservatively.

METHOD

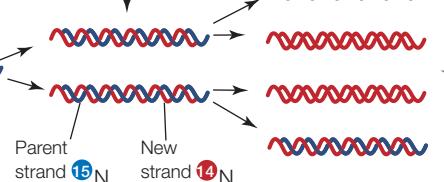


RESULTS



INTERPRETATION

Before the bacteria reproduce for the first time in the light medium (at 0 minutes), all DNA (parental) is heavy.



After two generations, half the DNA was intermediate and half was light; there was no heavy DNA.

CONCLUSION This pattern could only have been observed if each DNA molecule contains a template strand from the parental DNA; thus DNA replication is semiconservative.

FURTHER INVESTIGATION: If you continued this experiment for two more generations (as Meselson and Stahl actually did), what would be the composition (in terms of density) of the fourth generation DNA?

Go to yourBioPortal.com for original citations, discussions, and relevant links for all INVESTIGATING LIFE figures.

every 20 minutes. Meselson and Stahl collected some of the bacteria after each division and extracted DNA from the samples. They found that the density gradient was different in each bacterial generation:

- At the time of the transfer to the ^{14}N medium, the DNA was uniformly labeled with ^{15}N , and hence formed a single band corresponding with dense DNA.
- After one generation in the ^{14}N medium, when the DNA had been duplicated once, all the DNA was of intermediate density.
- After two generations, there were two equally large DNA bands: one of low density and one of intermediate density.
- In samples from subsequent generations, the proportion of low-density DNA increased steadily.

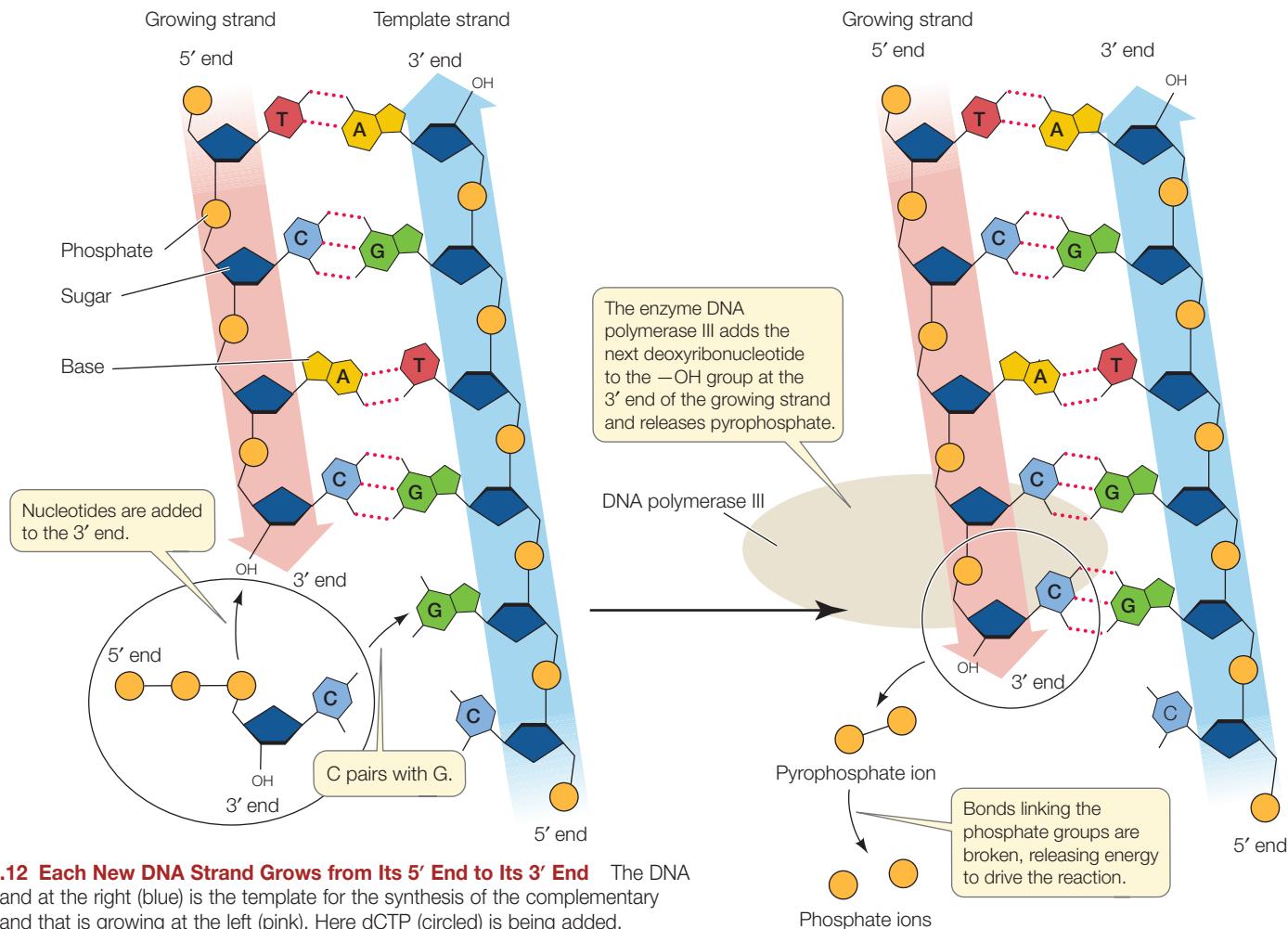
The results of this experiment can be explained only by the semiconservative model of DNA replication. In the first round of DNA replication in the ^{14}N medium, the strands of the double helix—both heavy with ^{15}N —separated. Each strand then acted as the template for a second strand, which contained only ^{14}N and hence was less dense. Each double helix then consisted of one ^{15}N strand and one ^{14}N strand, and was of intermediate

density. In the second replication, the ^{14}N -containing strands directed the synthesis of partners with ^{14}N , creating low-density DNA, and the ^{15}N strands formed new ^{14}N partners.

The crucial observation demonstrating the semiconservative model was that intermediate-density DNA (^{15}N – ^{14}N) appeared in the first generation and continued to appear in subsequent generations. With the other models, the results would have been quite different (see Figure 13.10):

- If conservative replication had occurred, the first generation would have had both high-density DNA (^{15}N – ^{15}N) and low-density DNA (^{14}N – ^{14}N), but no intermediate-density DNA.
- If dispersive replication had occurred, the density of the new DNA would have been intermediate, but DNA of this density would not continue to appear in subsequent generations.

Some scientists consider the Meselson–Stahl experiment to be one of the most elegant experiments ever performed by biologists, and it is an excellent example of the scientific method. It began with three hypotheses—the three models of DNA replication—and was designed so that the results could differentiate between them.



13.12 Each New DNA Strand Grows from Its 5' End to Its 3' End The DNA strand at the right (blue) is the template for the synthesis of the complementary strand that is growing at the left (pink). Here dCTP (circled) is being added.

There are two steps in DNA replication

Semiconservative DNA replication in the cell involves a number of different enzymes and other proteins. It takes place in two general steps:

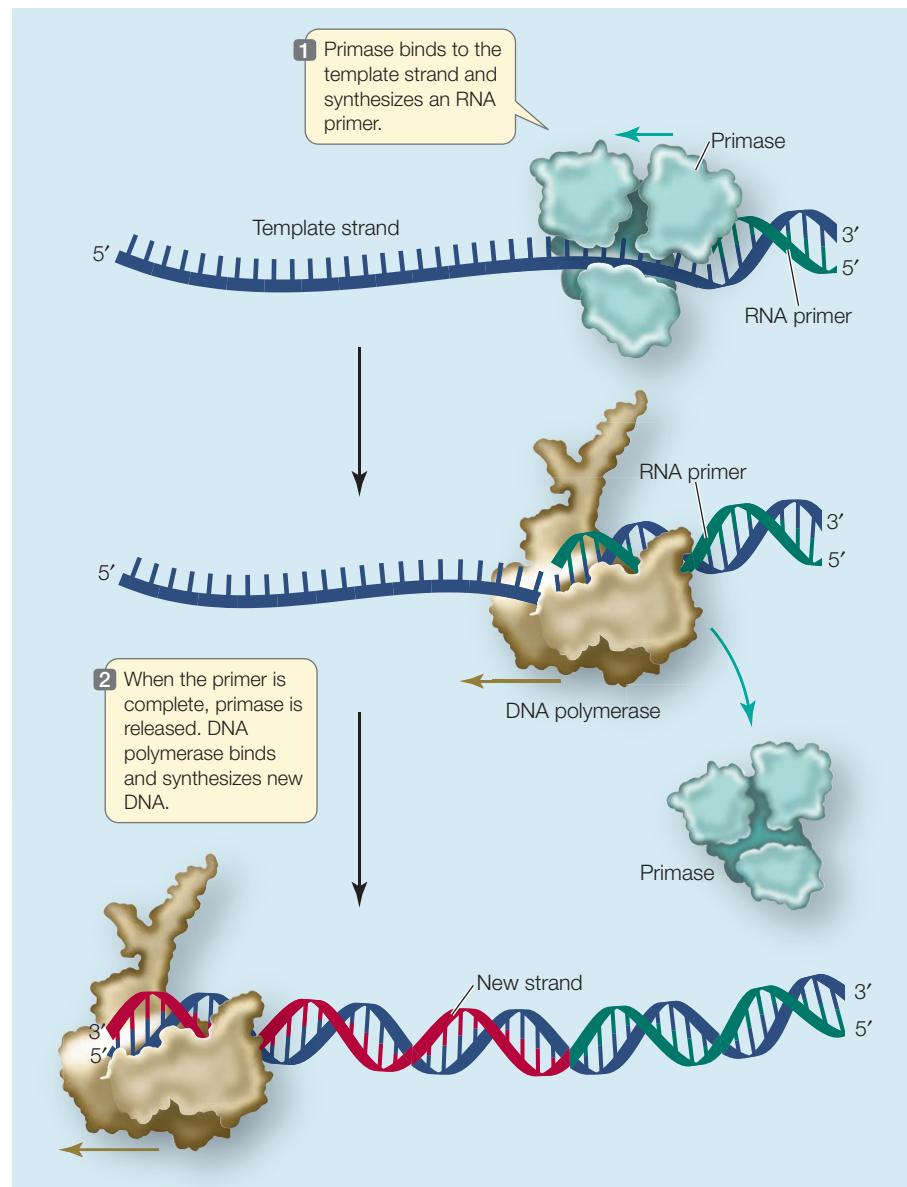
- The DNA double helix is unwound to separate the two template strands and make them available for new base pairing.
- As new nucleotides form complementary base pairs with template DNA, they are covalently linked together by phosphodiester bonds, forming a polymer whose base sequence is complementary to the bases in the template strand.

A key observation is that *nucleotides are added to the growing new strand at the 3' end*—the end at which the DNA strand has a free hydroxyl ($-OH$) group on the 3' carbon of its terminal deoxyribose (Figure 13.12). One of the three phosphate groups in a dNTP is attached to the 5' position of the sugar. The bonds linking the other two phosphate groups to the dNTP are broken, resulting in a monophosphate nucleotide, and releasing energy for the reaction.

DNA polymerases add nucleotides to the growing chain

DNA is replicated through the interaction of the template strand with a huge protein complex called the **replication complex**, which contains at least four proteins, including DNA polymerase. All chromosomes have at least one region called the **origin of replication (ori)**, to which the replication complex binds. Binding occurs when proteins in the complex recognize a specific DNA sequence within the origin of replication.

DNA REPLICATION BEGINS WITH A PRIMER A DNA polymerase elongates a polynucleotide strand by covalently linking new nucleotides to a previously existing strand. However, it cannot start this process without a short “starter” strand, called a **primer**. In DNA replication, the primer is usually a short single strand of RNA (Figure 13.13) but in some organisms it is DNA. This RNA primer strand is complementary to the DNA template, and is synthesized one nucleotide at a time by an enzyme called a **primase**. The DNA polymerase then adds nucleotides to the 3' end of the primer and continues until the replication of that section of DNA has been completed. Then the RNA primer is degraded, DNA is added in its place, and the resulting DNA fragments are connected by the action of

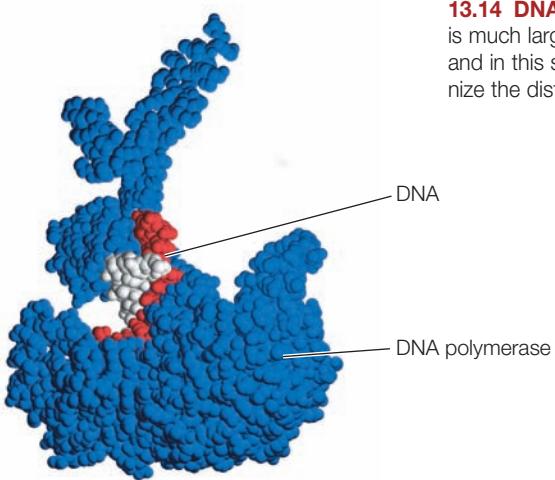


13.13 DNA Forms with a Primer DNA polymerases require a primer—a “starter” strand of DNA or RNA to which they can add new nucleotides.

other enzymes. When DNA replication is complete, each new strand consists only of DNA.

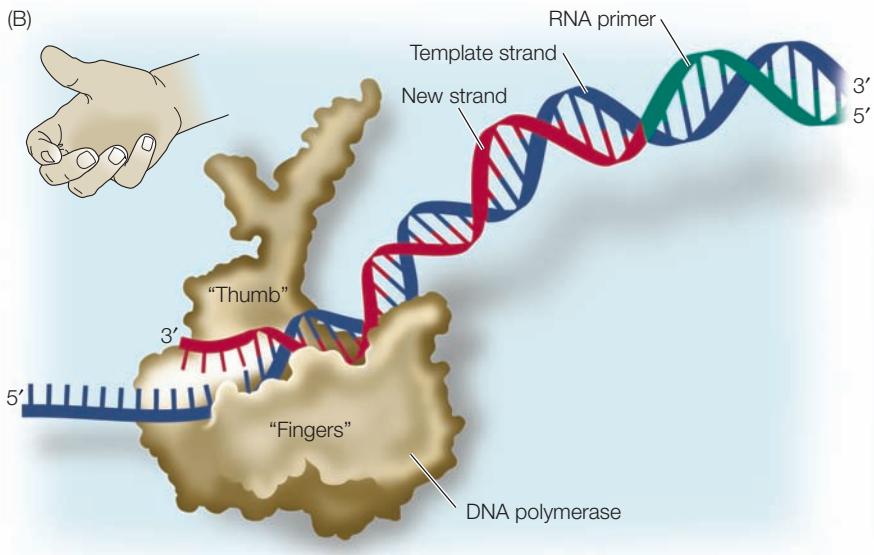
DNA POLYMERASES ARE LARGE DNA polymerases are much larger than their substrates, the dNTPs, and the template DNA, which is very thin (Figure 13.14A). Molecular models of the enzyme–substrate–template complex from bacteria show that the enzyme is shaped like an open right hand with a palm, a thumb, and fingers (Figure 13.14B). The palm holds the active site of the enzyme and brings together each substrate and the template. The finger regions rotate inward and have precise shapes that can recognize the different shapes of the four nucleotide bases.

(A)



13.14 DNA Polymerase Binds to the Template Strand (A) The DNA polymerase enzyme (blue) is much larger than the DNA molecule (red and white). (B) DNA polymerase is shaped like a hand, and in this side-on view, its “fingers” can be seen curling around the DNA. These “fingers” can recognize the distinctive shapes of the four bases.

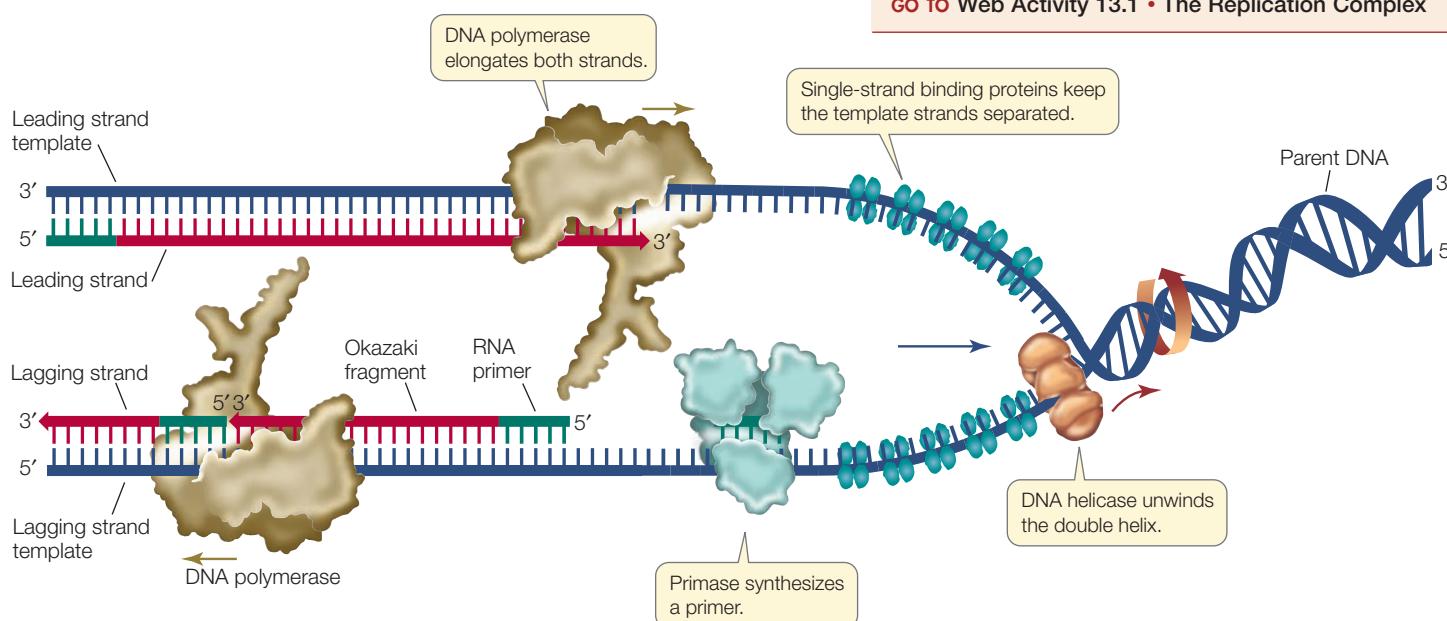
(B)



CELLS CONTAIN SEVERAL DIFFERENT DNA POLYMERASES Most cells contain more than one kind of DNA polymerase, but only one of them is responsible for chromosomal DNA replication. The others are involved in primer removal and DNA repair. Fifteen DNA polymerases have been identified in humans; the ones catalyzing most replication are DNA polymerases δ (delta) and ϵ (epsilon). In the bacterium *E. coli* there are five DNA polymerases; the one responsible for replication is DNA polymerase III.

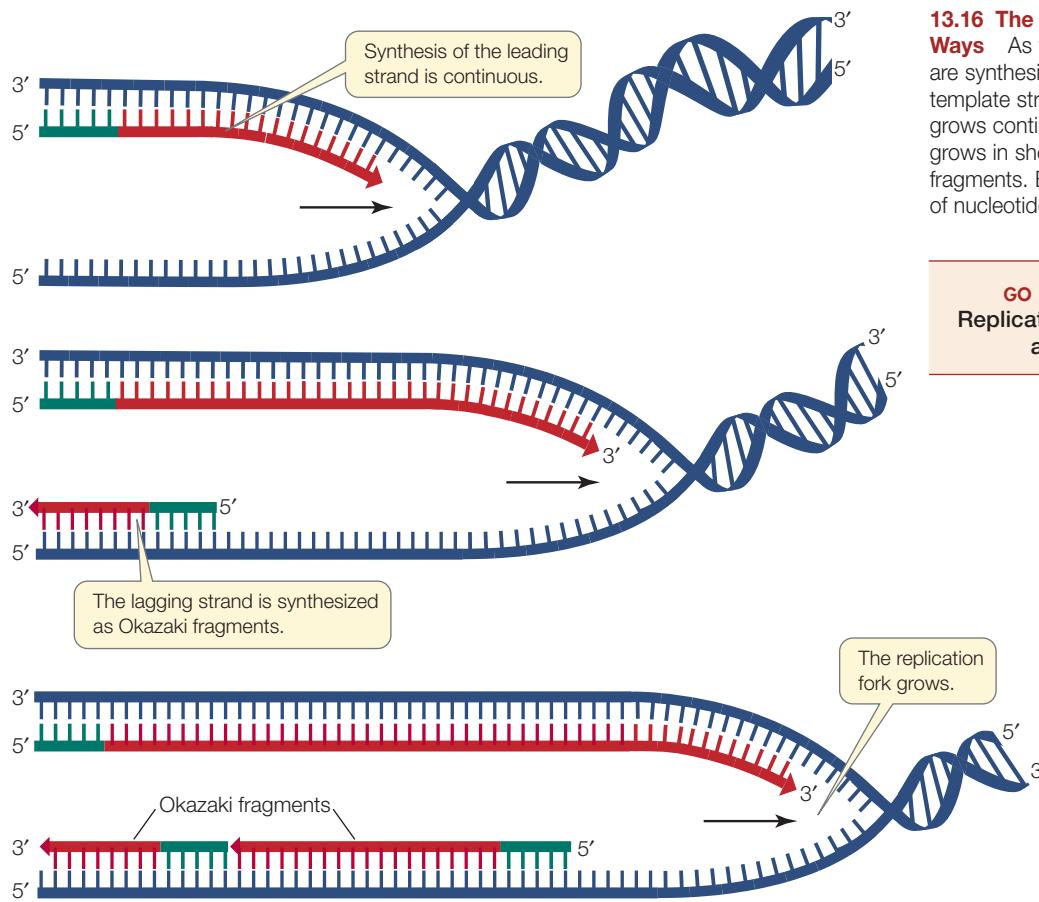
Many other proteins assist with DNA polymerization

Various other proteins play roles in other replication tasks; some of these are shown in **Figure 13.15**. The first event at the origin of replication is the localized unwinding and separation (denaturation) of the DNA strands. As we saw in Chapter 4, there are several forces that hold the two strands together, including hydrogen bonding and the hydrophobic interactions of the bases. An enzyme called **DNA helicase** uses energy from ATP hydrolysis to unwind and separate the strands, and spe-



yourBioPortal.com

GO TO Web Activity 13.1 • The Replication Complex



13.16 The Two New Strands Form in Different Ways

As the parent DNA unwinds, both new strands are synthesized in the 5'-to-3' direction, although their template strands are antiparallel. The leading strand grows continuously forward, but the lagging strand grows in short discontinuous stretches called Okazaki fragments. Eukaryotic Okazaki fragments are hundreds of nucleotides long, with gaps between them.

yourBioPortal.com

GO TO Animated Tutorial 13.3 • DNA Replication, Part 2: Coordination of Leading and Lagging Strand Synthesis

cial proteins called **single-strand binding proteins** bind to the unwound strands to keep them from reassociating into a double helix. This process makes each of the two template strands available for complementary base pairing.

THE TWO DNA STRANDS GROW DIFFERENTLY As Figure 13.15 shows, the DNA at the **replication fork**—the site(s) where DNA unwinds to expose the bases so that they can act as templates—opens up like a zipper in one direction. Study **Figure 13.16** and try to imagine what is happening over a short period of time. Remember that the two DNA strands are antiparallel; that is, the 3' end of one strand is paired with the 5' end of the other.

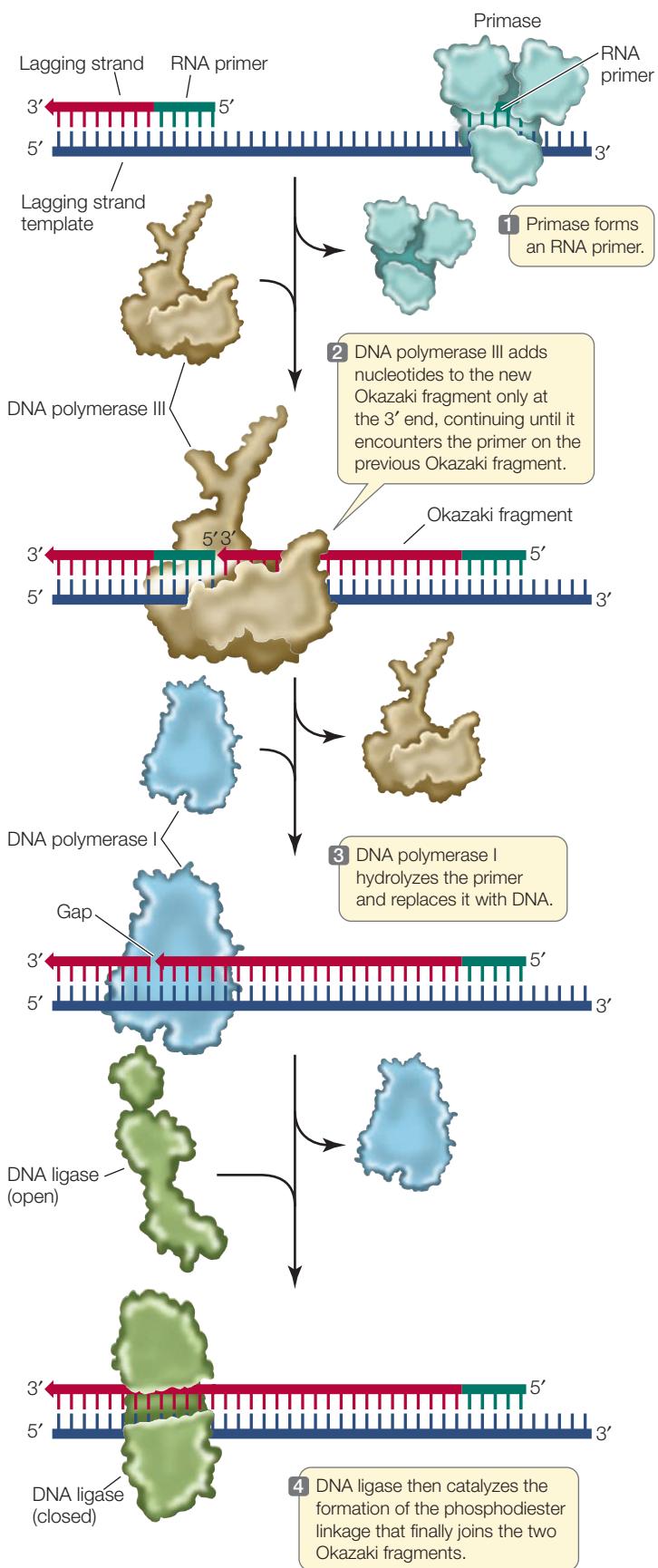
- One newly replicating strand (the **leading strand**) is oriented so that it can grow continuously at its 3' end as the fork opens up.
- The other new strand (the **lagging strand**) is oriented so that as the fork opens up, its exposed 3' end gets farther and farther away from the fork, and an unreplicated gap is formed. This gap would get bigger and bigger if there were not a special mechanism to overcome this problem.

Synthesis of the lagging strand requires the synthesis of relatively small, discontinuous stretches of sequence (100 to 200 nucleotides in eukaryotes; 1,000 to 2,000 nucleotides in prokaryotes). These discontinuous stretches are synthesized just as the

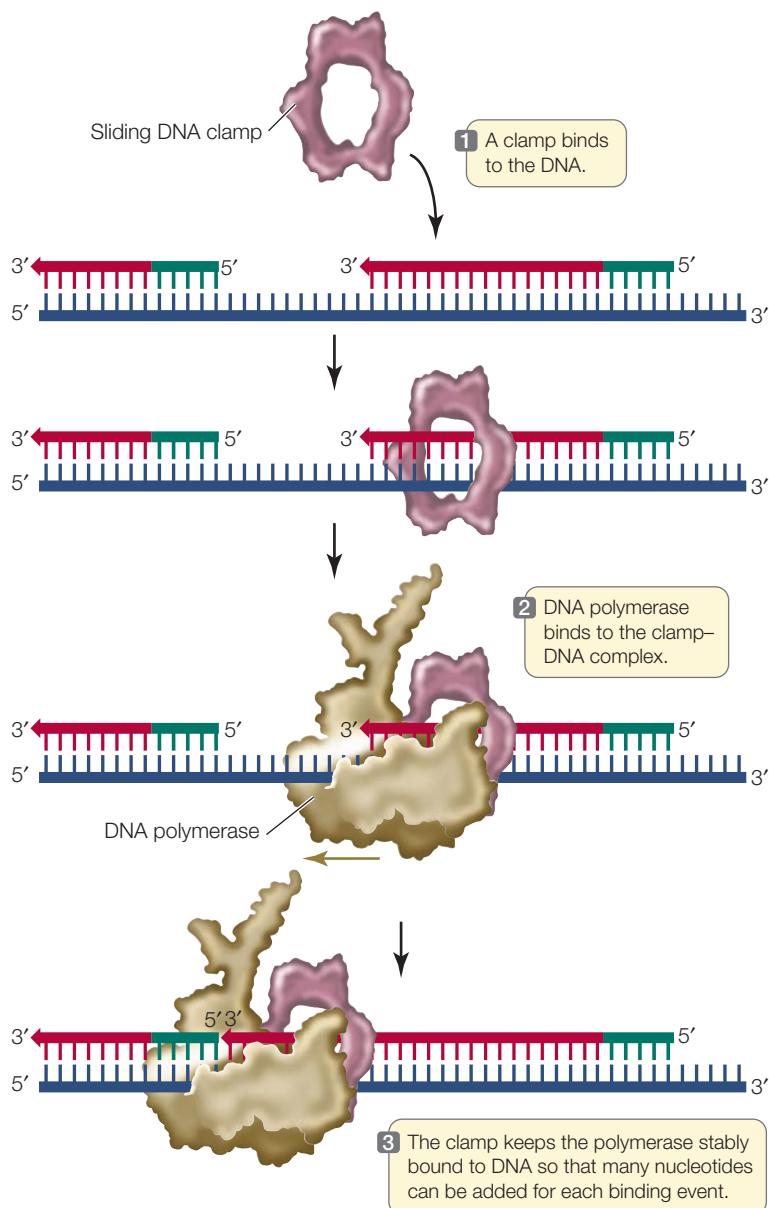
leading strand is, by the addition of new nucleotides one at a time to the 3' end of the new strand, but the synthesis of this new strand moves in the direction opposite to that in which the replication fork is moving. These stretches of new DNA are called **Okazaki fragments** (after their discoverer, the Japanese biochemist Reiji Okazaki). While the leading strand grows continuously “forward,” the lagging strand grows in shorter, “backward” stretches with gaps between them.

A single primer is needed for synthesis of the leading strand, but each Okazaki fragment requires its own primer to be synthesized by the primase. In bacteria, DNA polymerase III then synthesizes an Okazaki fragment by adding nucleotides to one primer until it reaches the primer of the previous fragment. At this point, DNA polymerase I (discovered by Arthur Kornberg) removes the old primer and replaces it with DNA. Left behind is a tiny nick—the final phosphodiester linkage between the adjacent Okazaki fragments is missing. The enzyme **DNA ligase** catalyzes the formation of that bond, linking the fragments and making the lagging strand whole (**Figure 13.17**).

Working together, DNA helicase, the two DNA polymerases, primase, DNA ligase, and the other proteins of the replication complex do the job of DNA synthesis with a speed and accuracy that are almost unimaginable. In *E. coli*, the replication complex makes new DNA at a rate in excess of 1,000 base pairs per second, committing errors in fewer than one base in a million.



13.17 The Lagging Strand Story In bacteria, DNA polymerase I and DNA ligase cooperate with DNA polymerase III to complete the complex task of synthesizing the lagging strand.



13.18 A Sliding DNA Clamp Increases the Efficiency of DNA Polymerization

Polymerization The clamp increases the efficiency of polymerization by keeping the enzyme bound to the substrate, so the enzyme does not have to repeatedly bind to template and substrate.

A SLIDING CLAMP INCREASES THE RATE OF DNA REPLICATION How do DNA polymerases work so fast? We saw in Section 8.3 that an enzyme catalyzes a chemical reaction:

substrate binds to enzyme → one product is formed → enzyme is released → cycle repeats

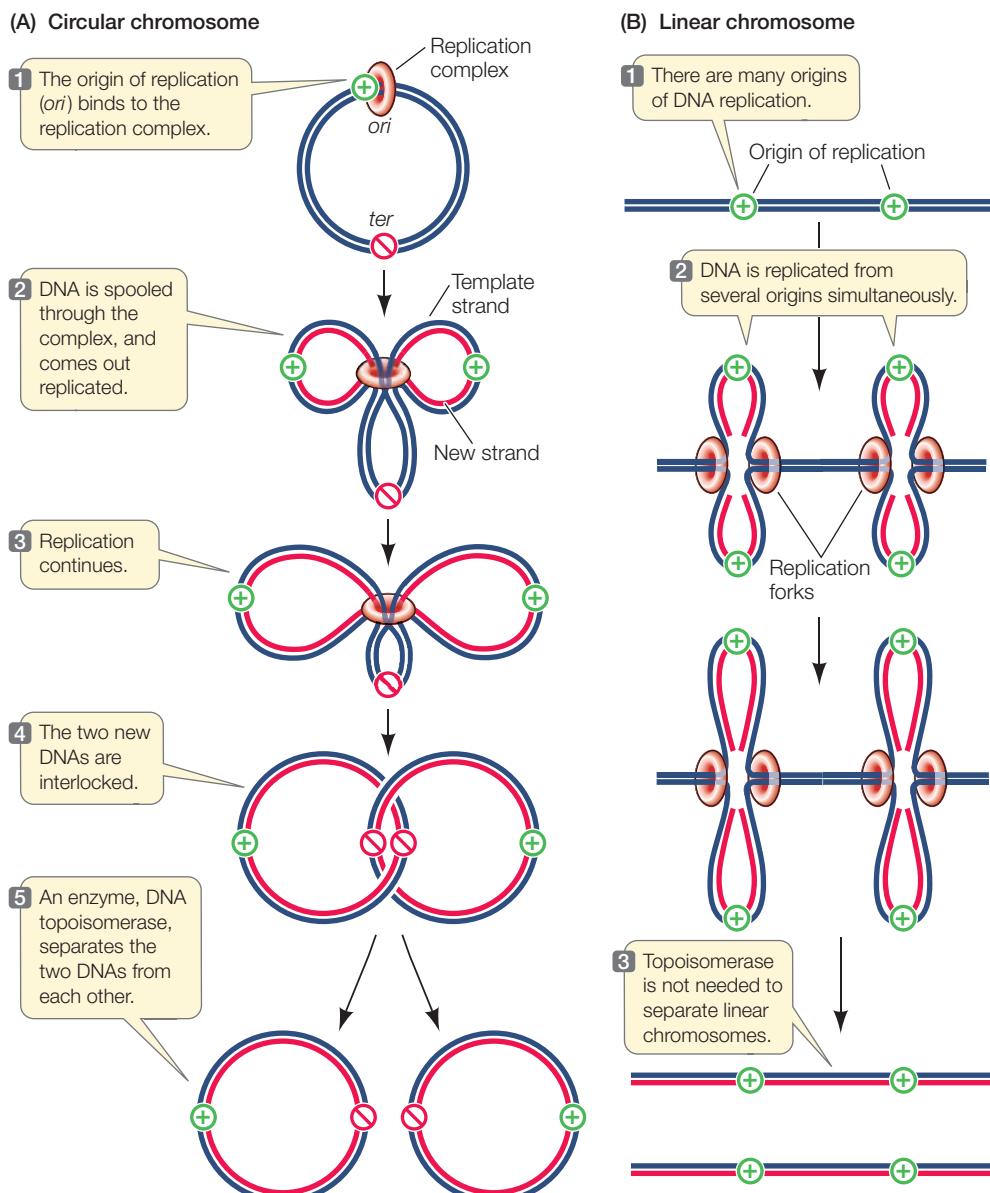
DNA replication would not proceed as rapidly as it does if it went through such a cycle for each nucleotide. Instead, DNA polymerases are **processive**—that is, they catalyze many polymerizations each time they bind to a DNA molecule:

substrates bind to one enzyme → many products are formed → enzyme is released → cycle repeats

The newly replicated strand is stabilized by a **sliding DNA clamp**, which is shaped like a screw cap on a bottle (Figure 13.18). This protein has multiple identical subunits assembled into a doughnut shape. The doughnut's "hole" is just large enough to en-

circle the DNA double helix, along with a single layer of water molecules for lubrication. The clamp binds to the DNA polymerase–DNA complex, keeping the enzyme and the DNA associated tightly with each other. If the clamp is absent, DNA polymerase dissociates from DNA after 20–100 polymerizations. With the clamp, it can polymerize up to 50,000 nucleotides before it detaches.

PCNA IS THE MAESTRO OF THE REPLICATION FORK In mammals, the sliding clamp was first recognized in rapidly dividing cells and is called **proliferating cell nuclear antigen (PCNA)**. PCNA does more than just keep the DNA polymerase bound to the DNA; it also helps to orient the polymerase for binding to the substrates. Furthermore, PCNA has binding sites for many other proteins, including chromosome structural proteins, DNA ligase, DNA methylation enzymes (see Section 16.4) and enzymes involved in DNA repair (see below). It also removes the prereplication complex from *ori*, ensuring that replication only happens once per cell cycle. For all that it does, PCNA has been called the “maestro of the replication fork.”



DNA IS THREADED THROUGH A REPLICATION COMPLEX Until recently, DNA replication was always depicted to look like a locomotive (the replication complex) moving along a railroad track (the DNA). While this does occur in some organisms, most commonly in eukaryotes the replication complex seems to be stationary, attached to chromatin structures, and it is the DNA that moves, essentially threading through the complex as single strands and emerging as double strands).

SMALL CIRCULAR CHROMOSOMES REPLICATE FROM A SINGLE ORIGIN Small circular chromosomes, such as those of bacteria (consisting of 1–4 million base pairs), have a single origin of replication. Two replication forks form at this *ori*, and as the DNA moves through the replication complex, the replication forks extend around the circle (Figure 13.19A). Two interlocking circular DNA molecules are formed, and they are separated by an enzyme called **DNA topoisomerase**. As we mentioned above, DNA polymerases are very fast. In *E. coli*, replication can be as fast as 1,000 bases per second, and it takes 20–40 minutes to replicate the bacterium’s 4.7 million base pairs.

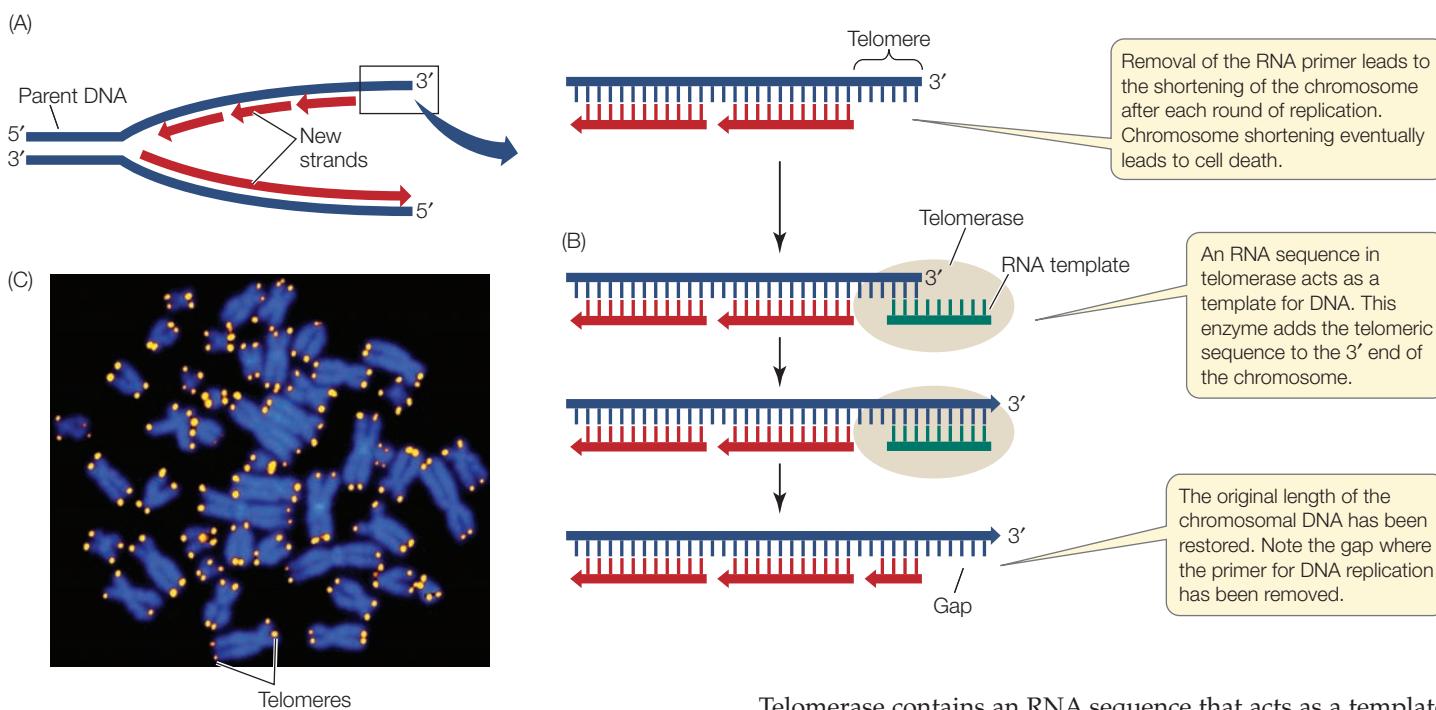
LARGE LINEAR CHROMOSOMES HAVE MANY ORIGINS Human DNA polymerases are slower than those of *E. coli*, and can replicate DNA at a rate of about 50 bases per second. Human chromosomes are much larger than those of bacteria (about 80 million base pairs) and linear. Large linear chromosomes such as those of humans contain hundreds of origins of replication. Numerous replication complexes bind to these sites at the same time and catalyze simultaneous replication. Thus there are many replication forks in eukaryotic DNA (Figure 13.19B).

Telomeres are not fully replicated and are prone to repair

As we have just seen, replication of the lagging strand occurs by the addition of Okazaki fragments to RNA primers. When the terminal RNA primer is removed, no DNA can be synthesized to replace it because there is no 3' end to extend (Figure 13.20A). So the new chromosome has a bit of single-stranded DNA at each end. This situation activates a mechanism for cutting off the single-stranded region, along with some of the

13.19 Replication of Small Circular and Large Linear Chromosomes

(A) Small circular chromosomes, typical of prokaryotes, have a single origin (*ori*) and terminus (*ter*) of replication. (B) Larger linear chromosomes, typical of nuclear DNA in eukaryotes, have many origins of replication.



13.20 Telomeres and Telomerase (A) Removal of the RNA primer at the 3' end of the template for the lagging strand leaves a region of DNA—the telomere—unreplicated. (B) In continuously dividing cells, the enzyme telomerase binds to the 3' end and extends the lagging strand of DNA, so the chromosome does not get shorter. (C) Bright fluorescent staining marks the telomeric regions on these blue-stained human chromosomes.

intact double-stranded DNA. Thus the chromosome becomes slightly shorter with each cell division.

There is another, more serious problem at the ends of chromosomes, and that is simply that they are ends! In Section 11.2 we described checkpoints in the cell cycle for the integrity of DNA. At one of the checkpoints, the DNA is examined for DNA breaks (due to radiation, etc.) and DNA repair is initiated if breaks are found. This involves joining the breaks via a combination of DNA synthesis and DNA ligase activity. This system might recognize the ends of chromosomes as breaks and join two chromosomes together. This would create havoc with genomic integrity.

In many eukaryotes, there are repetitive sequences at the ends of chromosomes called **telomeres**. In humans, the telomere sequence is TTAGGG, and it is repeated about 2,500 times. These repeats bind special proteins that prevent the DNA repair system from recognizing the ends as breaks. In addition, the repeats may form loops that have a similar protective role. So the telomere acts like the plastic tip of shoelaces to prevent fraying.

Each human chromosome can lose 50–200 base pairs of telomeric DNA after each round of DNA replication and cell division. After 20–30 divisions, the chromosomes are unable to participate in cell division, and the cell dies. This phenomenon explains, in part, why many cell lineages do not last the entire lifetime of the organism: their telomeres are lost. Yet continuously dividing cells, such as bone marrow stem cells and gamete-producing cells, maintain their telomeric DNA. An enzyme, appropriately called **telomerase**, catalyzes the addition of any lost telomeric sequences in these cells (**Figure 13.20B**).

Telomerase contains an RNA sequence that acts as a template for the telomeric DNA repeat sequence.

Telomerase is expressed in more than 90 percent of human cancers, and may be an important factor in the ability of cancer cells to divide continuously. Since most normal cells do not have this ability, telomerase is an attractive target for drugs designed to attack tumors specifically.

There is also interest in telomerase and aging. When a gene expressing high levels of telomerase is added to human cells in culture, their telomeres do not shorten. Instead of living 20–30 cell generations and then dying, the cells become immortal. It remains to be seen how this finding relates to the aging of a whole organism.

13.3 RECAP

Meselson and Stahl showed that DNA replication is semiconservative: each parent DNA strand serves as a template for a new strand. A complex of proteins, most notably DNA polymerases, is involved in replication. New DNA is polymerized in one direction only, and since the two strands are antiparallel, one strand is made continuously and the other is synthesized in short Okazaki fragments that are eventually joined.

- How did the Meselson–Stahl experiment differentiate between the three models for DNA replication? See pp. 276–278 and Figures 13.10 and 13.11
- What are the five enzymes needed for DNA replication and what are their roles? See pp. 279–283 and Figures 13.13–13.17
- How is the leading strand of DNA replicated continuously while the lagging strand must be replicated in fragments? See p. 281 and Figure 13.16

The complex process of DNA replication is amazingly accurate, but it is not perfect. What happens when things go wrong?

13.4 How Are Errors in DNA Repaired?

DNA must be accurately replicated and faithfully maintained. The price of failure can be great; the accurate transmission of genetic information is essential for the functioning and even the life of a single cell or multicellular organism. Yet the replication of DNA is not perfectly accurate, and the DNA of nondividing cells is subject to damage by natural chemical alterations and by environmental agents. In the face of these threats, how has life gone on so long?

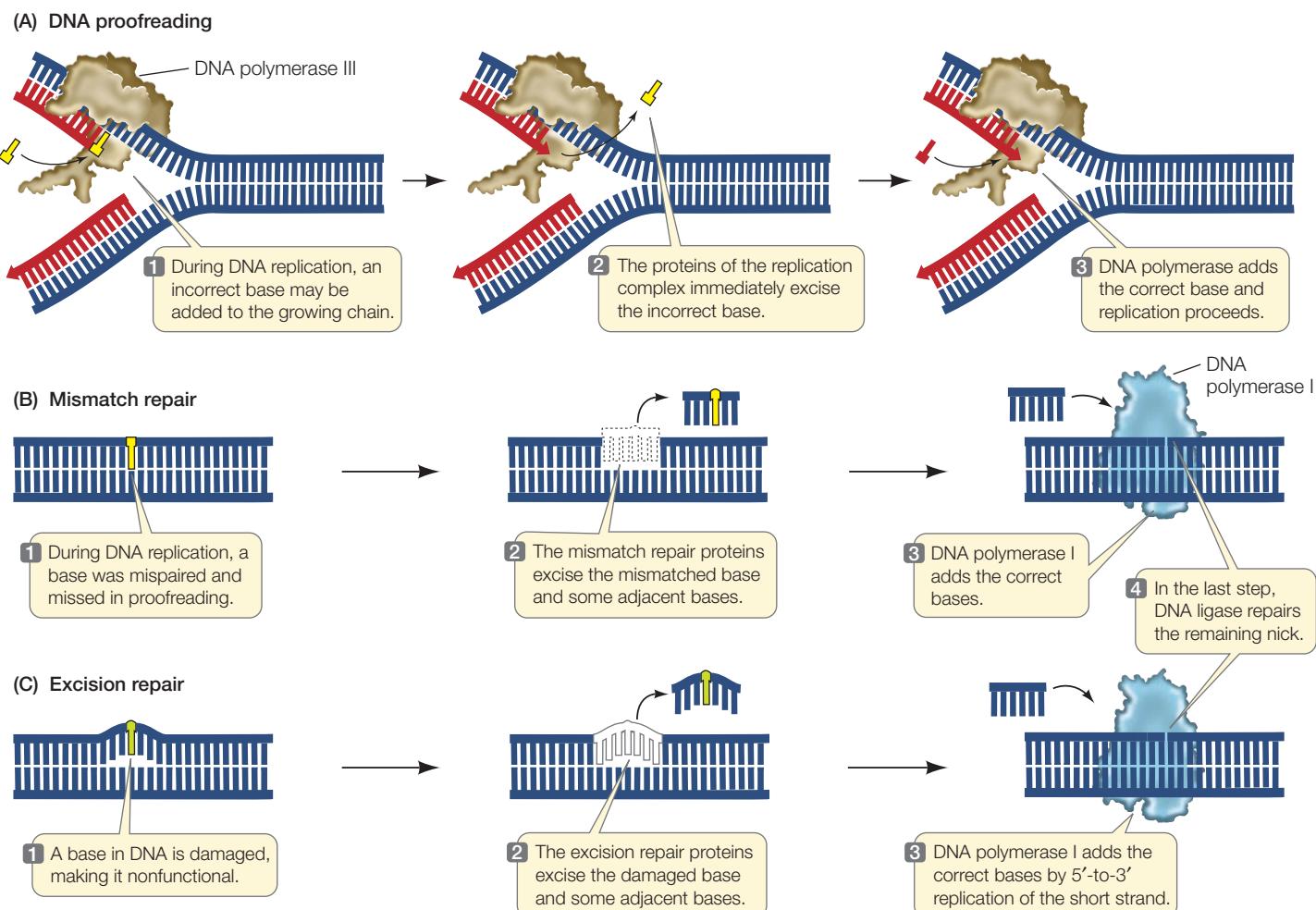
DNA repair mechanisms help to preserve life. DNA polymerases initially make significant numbers of mistakes in assembling polynucleotide strands. Without DNA repair, the observed error rate of one for every 10^5 bases replicated would result in about 60,000 mutations every time a human cell divided. Fortunately, our cells can repair damaged nucleotides and DNA replication errors, so that very few errors end up in the replicated DNA. Cells have at least three DNA repair mechanisms at their disposal:

- A **proofreading** mechanism corrects errors in replication as DNA polymerase makes them.

- A **mismatch repair** mechanism scans DNA immediately after it has been replicated and corrects any base-pairing mismatches.
- An **excision repair** mechanism removes abnormal bases that have formed because of chemical damage and replaces them with functional bases.

Most DNA polymerases perform a **proofreading** function each time they introduce a new nucleotide into a growing DNA strand (Figure 13.21A). When a DNA polymerase recognizes a mispairing of bases, it removes the improperly introduced nucleotide and tries again. (Other proteins in the replication complex also play roles in proofreading.) The error rate for this process is only about 1 in 10,000 repaired base pairs, and it lowers the overall error rate for replication to about one error in every 10^{10} bases replicated.

After the DNA has been replicated, a second set of proteins surveys the newly replicated molecule and looks for mismatched base pairs that were missed in proofreading (Figure



13.21 DNA Repair Mechanisms The proteins of the replication complex function in DNA repair mechanisms, reducing the rate of errors in the replicated DNA. Another mechanism (excision repair) repairs damage to existing DNA molecules.

13.21B). For example, this mismatch repair mechanism might detect an A-C base pair instead of an A-T pair. But how does the repair mechanism “know” whether the A-C pair should be repaired by removing the C and replacing it with T or by removing the A and replacing it with G?

The mismatch repair mechanism can detect the “wrong” base because a DNA strand is chemically modified some time after replication. In prokaryotes, methyl groups ($-\text{CH}_3$) are added to some adenines. In eukaryotes, cytosine bases are methylated. Immediately after replication, methylation has not yet occurred on the newly replicated strand, so the new strand is “marked” (distinguished by being unmethylated) as the one in which errors should be corrected.

When mismatch repair fails, DNA sequences are altered. One form of colon cancer arises in part from a failure of mismatch repair.

DNA molecules can also be damaged during the life of a cell (for example, when it is in G_1). High-energy radiation, chemicals from the environment, and random spontaneous chemical reactions can all damage DNA. **Excision repair** mechanisms deal with these kinds of damage (Figure 13.21C). Individuals who suffer from a condition known as xeroderma pigmentosum lack an excision repair mechanism that normally corrects the damage caused by ultraviolet radiation. They can develop skin cancers after even a brief exposure to sunlight.

13.4 RECAP

DNA replication is not perfect; in addition, DNA may be naturally altered or damaged. Repair mechanisms exist that detect and repair mismatched or damaged DNA.

- Explain the roles of DNA proofreading, mismatch repair, and excision repair. See Figure 13.21

Understanding how DNA is replicated and repaired has allowed scientists to develop techniques for studying genes. We'll look at just one of those techniques next.

13.5 How Does the Polymerase Chain Reaction Amplify DNA?

The principles underlying DNA replication in cells have been used to develop an important laboratory technique that has been vital in analyzing genes and genomes. This technique allows researchers to make multiple copies of short DNA sequences.

The polymerase chain reaction makes multiple copies of DNA sequences

In order to study DNA and perform genetic manipulations, it is necessary to make multiple copies of a DNA sequence. This is necessary because the amount of DNA isolated from a biological sample is often too small to work with. The **polymerase**

chain reaction (PCR) technique essentially automates this replication process by copying a short region of DNA many times in a test tube. This process is referred to as *DNA amplification*.

The PCR reaction mixture contains:

- a sample of double stranded DNA from a biological sample, to act as the template,
- two short, artificially synthesized primers that are complementary to the ends of the sequence to be amplified,
- the four dNTPs (dATP, dTTP, dCTP and dGTP),
- a DNA polymerase that can tolerate high temperatures without becoming degraded, and
- salts and a buffer to maintain a near-neutral pH.

The PCR amplification is a cyclic process in which a sequence of steps is repeated over and over again (Figure 13.22):

- The first step involves heating the reaction to near boiling point, to separate (*denature*) the two strands of the DNA template.
- The reaction is then cooled to allow the primers to bind (or *anneal*) to the template strands.
- Next, the reaction is warmed to an optimum temperature for the DNA polymerase to catalyze the production of the complementary new strands.

A single cycle takes a few minutes to produce two copies of the target DNA sequence, leaving the new DNA in the double-stranded state. Repeating the cycle many times leads to an exponential increase in the number of copies of the DNA sequence.

The PCR technique requires that the base sequences at the 3' end of each strand of the target DNA sequence be known, so that complementary primers, usually 15–30 bases long, can be made in the laboratory. Because of the uniqueness of DNA sequences, a pair of primers this length will usually bind to only a single region of DNA in an organism's genome. This specificity, despite the incredible diversity of DNA sequences, is a key to the power of PCR.

One initial problem with PCR was its temperature requirements. To denature the DNA, it must be heated to more than 90°C—a temperature that destroys most DNA polymerases. The PCR technique would not be practical if new polymerase had to be added after denaturation in each cycle.

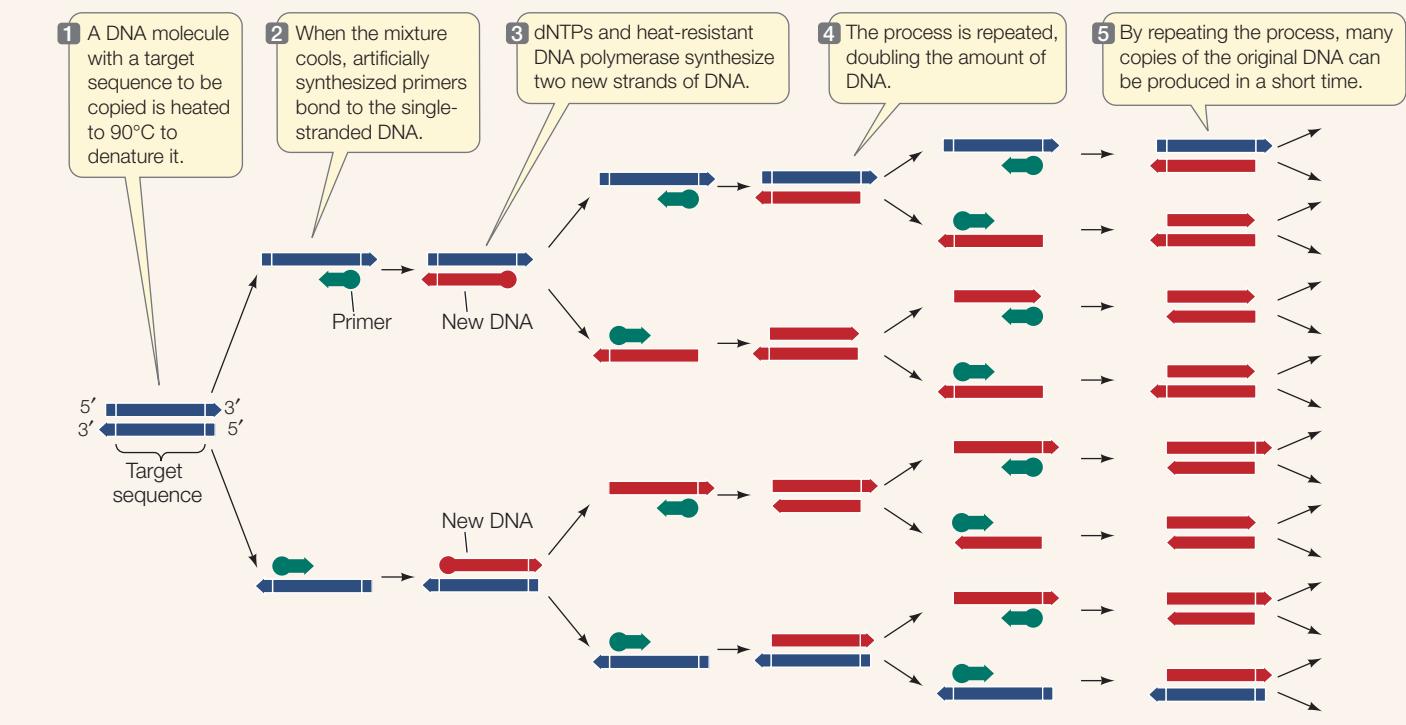
This problem was solved by nature: in the hot springs at Yellowstone National Park, as well as in other high-temperature locations, there lives a bacterium called, appropriately, *Thermus aquaticus* (“hot water”). The means by which this organism survives temperatures of up to 95°C was investigated by Thomas Brock and his colleagues at the University of Wisconsin, Madison. They discovered that *T. aquaticus* has an entire metabolic machinery that is heat-resistant, including a DNA polymerase that does not denature at these high temperatures.

Scientists pondering the problem of copying DNA by PCR read Brock's basic research articles and got a clever idea: why not use *T. aquaticus* DNA polymerase in the PCR technique? It

TOOLS FOR INVESTIGATING LIFE

13.22 The Polymerase Chain Reaction

The steps in this cyclic process are repeated many times to produce multiple identical copies of a DNA fragment. This makes enough DNA for chemical analysis and genetic manipulations.



could withstand the 90°C denaturation temperature and would not have to be added during each cycle. The idea worked, and it earned biochemist Kary Mullis a Nobel prize. PCR has had an enormous impact on genetic research. Some of its most striking applications will be described in Chapters 15–18. These applications range from amplifying DNA in order to identify an individual person or organism, to detection of diseases.

13.5 RECAP

Knowledge of the mechanisms of DNA replication led to the development of a technique for making multiple copies of DNA sequences.

- What is the role of primers in PCR? See pp. 286 and Figure 13.22

CHAPTER SUMMARY

13.1 What Is the Evidence that the Gene Is DNA?

- Griffith's experiments in the 1920s demonstrated that some substance in cells—then called a **transforming principle**—can cause heritable changes in other cells. Review Figure 13.1
- The location and quantity of DNA in the cell suggested that DNA might be the genetic material. Avery, MacLeod, and McCarty isolated the transforming principle from bacteria and identified it as DNA. Review Figure 13.2
- The Hershey–Chase experiment established conclusively that DNA (and not protein) is the genetic material, by tracing the

DNA of radioactively labeled viruses, with which they infected bacterial cells. Review Figure 13.4

- Genetic transformation of eukaryotic cells is called **transfection**. Transformation and transfection can be studied with the aid of a **marker** gene that confers a known and observable phenotype. Review Figure 13.5

13.2 What Is the Structure of DNA?

- Chargaff's rule states that the amount of **adenine** in DNA is equal to the amount of **thymine**, and that the amount of **guanine** is equal to the amount of **cytosine**; thus the total

abundance of purines ($A + G$) equals the total abundance of pyrimidines ($T + C$).

- X-ray crystallography showed that the DNA molecule is **helical**. Watson and Crick proposed that DNA is a double-stranded helix in which the strands are **antiparallel**. **Review Figure 13.8**
- **Complementary base pairing** between A and T and between G and C accounts for Chargaff's rule. The bases are held together by hydrogen bonding. **Review Figure 13.9**

13.3 How Is DNA Replicated?

SEE ANIMATED TUTORIAL 13.1

- Meselson and Stahl showed the replication of DNA to be **semi-conservative**. Each parent strand acts as a **template** for the synthesis of a new strand; thus the two replicated DNA molecules each contain one parent strand and one newly synthesized strand. **Review Figure 13.11, ANIMATED TUTORIAL 13.2**
- In DNA replication, the enzyme **DNA polymerase** catalyzes the addition of nucleotides to the 3' end of each strand. Which nucleotides are added is determined by complementary base pairing with the template strand. **Review Figure 13.12**
- The **replication complex** is a huge protein complex that attaches to the chromosome at the **origin of replication (ori)**.
- Replication proceeds from the origin of replication on both strands in the 5'-to-3' direction, forming a **replication fork**.
- **Primase** catalyzes the synthesis of a short RNA **primer** to which nucleotides are added by DNA polymerase. **Review Figure 13.13**
- Many proteins assist in DNA replication. **DNA helicase** separates the strands, and **single-strand binding proteins** keep the strands from reassociating. **Review Figure 13.13, WEB ACTIVITY 13.1**

- The **leading strand** is synthesized continuously and the **lagging strand** in pieces called **Okazaki fragments**. The fragments are joined together by **DNA ligase**. **Review Figures 13.16 and 13.17, ANIMATED TUTORIAL 13.3**
- The speed with which DNA polymerization proceeds is attributed to the **processive** nature of DNA polymerases, which can catalyze many polymerizations at a time. A **sliding DNA clamp** helps ensure the stability of this process. **Review Figure 13.18**
- In prokaryotes, two interlocking circular DNA molecules are formed; they are separated by an enzyme called **DNA topoisomerase**. **Review Figure 13.19**
- In eukaryotes, DNA replication leaves a short, unreplicated sequence, the **telomere**, at the 3' end of the chromosome. Unless the enzyme **telomerase** is present, the sequence is removed. After multiple cell cycles, the telomeres shorten, leading to chromosome instability and cell death. **Review Figure 13.20**

13.4 How Are Errors in DNA repaired?

- DNA polymerases make about one error in 10^5 bases replicated. DNA is also subject to natural alterations and chemical damage. DNA can be repaired by three different mechanisms: **proof-reading**, **mismatch repair**, and **excision repair**. **Review Figure 13.21**

13.5 How Does the Polymerase Chain Reaction Amplify DNA?

- The **polymerase chain reaction** technique uses DNA polymerase to make multiple copies of DNA in the laboratory. **Review Figure 13.22**

SELF-QUIZ

1. Griffith's studies of *Streptococcus pneumoniae*
 - showed that DNA is the genetic material of bacteria.
 - showed that DNA is the genetic material of bacteriophage.
 - demonstrated the phenomenon of bacterial transformation.
 - proved that prokaryotes reproduce sexually.
 - proved that protein is not the genetic material.
2. In the Hershey-Chase experiment,
 - DNA from parent bacteriophage appeared in progeny bacteriophage.
 - most of the phage DNA never entered the bacteria.
 - more than three-fourths of the phage protein appeared in progeny phage.
 - DNA was labeled with radioactive sulfur.
 - DNA formed the coat of the bacteriophage.
3. Which statement about complementary base pairing is *not* true?
 - Complementary base pairing plays a role in DNA replication.
 - In DNA, T pairs with A.
 - Purines pair with purines, and pyrimidines pair with pyrimidines.
 - In DNA, C pairs with G.
 - The base pairs are of equal length.
4. In semiconservative replication of DNA,
 - the original double helix remains intact and a new double helix forms.
 - the strands of the double helix separate and act as templates for new strands.
 - polymerization is catalyzed by RNA polymerase.
 - polymerization is catalyzed by a double-helical enzyme.
 - DNA is synthesized from amino acids.
5. Which of the following does *not* occur during DNA replication?
 - Unwinding of the parent double helix
 - Formation of short pieces that are connected by DNA ligase
 - Complementary base pairing
 - Use of a primer
 - Polymerization in the 3'-to-5' direction
6. The primer used for DNA replication
 - is a short strand of RNA added to the 3' end.
 - is needed only once on a leading strand.
 - remains on the DNA after replication.
 - ensures that there will be a free 5' end to which nucleotides can be added.
 - is added to only one of the two template strands.

7. One strand of DNA has the sequence 5'-ATTCCG-3' The complementary strand for this is
 - a. 5'-TAAGGC-3'
 - b. 5'-ATTCCG-3'
 - c. 5'-ACCTTA-3'
 - d. 5'-CGGAAT-3'
 - e. 5'-GCCTTA-3'
8. The role of DNA ligase in DNA replication is to
 - a. add more nucleotides to the growing strand one at a time.
 - b. open up the two DNA strands to expose template strands.
 - c. ligate base to sugar to phosphate in a nucleotide.
 - d. bond Okazaki fragments to one another.
 - e. remove incorrectly paired bases.
9. The polymerase chain reaction
 - a. is a method for sequencing DNA.
 - b. is used to transcribe specific genes.
 - c. amplifies specific DNA sequences.
 - d. does not require DNA replication primers.
 - e. uses a DNA polymerase that denatures at 55°C.
10. What is the correct order for the following events in excision repair of DNA? (1) DNA polymerase I adds correct bases by 5' to 3' replication; (2) damaged bases are recognized; (3) DNA ligase seals the new strand to existing DNA; (4) part of a single strand is excised.
 - a. 1, 2, 3, 4
 - b. 2, 1, 3, 4
 - c. 2, 4, 1, 3
 - d. 3, 4, 2, 1
 - e. 4, 2, 3, 1

FOR DISCUSSION

1. Suppose that Meselson and Stahl had continued their experiment on DNA replication for another ten bacterial generations. Would there still have been any ^{14}N - ^{15}N hybrid DNA present? Would it still have appeared in the centrifuge tube? Explain.
2. If DNA replication were conservative rather than semiconservative, what results would Meselson and Stahl have observed? Draw a diagram of the results using the conventions of Figure 13.10.
3. Using the following information, calculate the number of origins of DNA replication on a human chromosome: DNA polymerase adds nucleotides at 3,000 base pairs per minute in one direction; replication is bidirectional; S phase lasts 300 minutes; there are 120 million base pairs per chromosome. In a typical chromosome 3 μm long, how many origins are there per μm ?
4. The drug dideoxycytidine, used to treat certain viral infections, is a nucleotide made with 2',3'-dideoxyribose. This sugar lacks —OH groups at both the 2' and the 3' positions. Explain why this drug stops the growth of a DNA chain when added to DNA.

ADDITIONAL INVESTIGATION

Outline a series of experiments using radioactive isotopes (such as ^{32}P and ^{35}S) to show that it is bacterial DNA and not bacterial

protein that enters the host cell and is responsible for bacterial transformation.

WORKING WITH DATA (GO TO yourBioPortal.com)

The Hershey-Chase Experiment The experiments in which labeled bacteriophage were used to infect host *E. coli* cells were key evidence for the identification of DNA as the gene (Figure 13.4). In this exercise, you will analyze the data that Hershey and Chase obtained, as well as important control experiments that ruled out protein and pointed to DNA as the gene.

The Meselson-Stahl Experiment Because of its elegant simplicity, this experiment has been called one of the most beautiful in the history of biology (Figure 13.11). In this real-world exercise, you will examine the experimental protocol and make calculations based on the actual centrifuge photographs that the experimenters obtained.

14

From DNA to Protein: Gene Expression

An unexpected wedding gift

The wedding and honeymoon began spectacularly. Andrew Speaker, an Atlanta lawyer, and law student Sarah Cooksey began their honeymoon in Rome. Days later, they got shocking news from the U.S. Centers for Disease Control and Prevention: Andrew had drug-resistant tuberculosis (TB) and would have to be quarantined to prevent him from spreading the disease to others.

Several months before, Speaker had gone to see his physician, complaining of a sore rib. The doctor ordered an X-ray and saw some fluid in Speaker's lungs. Suspicious, the physician sent samples of the fluid to a lab, which confirmed the diagnosis of TB. Moreover, the TB appeared to be resistant to several drugs.

Before the nineteenth-century German microbiologist Robert Koch identified the bacterium *Mycobacterium*

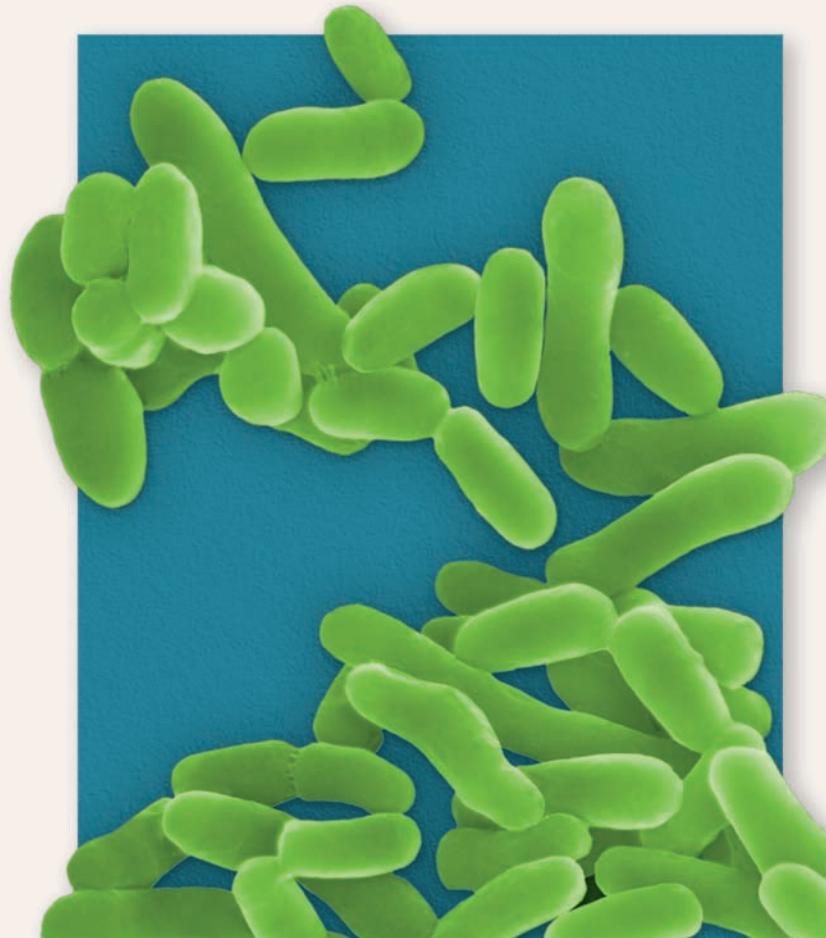
tuberculosis as its causative agent, TB was known as consumption. What started as a bloody cough with fever, chills, and night sweats would usually progress to other organs, including the nervous system. Death was almost inevitable. Today tuberculosis cases are fairly rare in the United States and Europe. Worldwide, however, there are more than 8 million cases and 1.6 million deaths annually, and TB remains the scourge it was in the nineteenth century. Speaker probably picked up the bacterium on his travels and it hid in his tissues (possibly for years) before flaring up at the time of his wedding.

Two drugs are used as the first approach to treating TB. One of them, isoniazid, is activated inside the bacterial cell, and the activated form blocks an enzyme essential for the assembly of the bacterial cell walls. Without functional cell walls, new bacterial cells cannot survive. The second drug, rifampin, binds to a part of the enzyme RNA polymerase that is necessary for gene expression. Without the appropriate expression of its genes, the bacterium soon dies.

In both cases, the targets of the antibiotics are proteins, each encoded by a gene (a sequence of DNA). Mutations in these genes can lead to altered amino acid sequences in the proteins, so that they no longer have three-dimensional structures that bind to the antibiotics. Unfortunately, these altered genes can be transferred from one bacterium to another, so a single *M. tuberculosis* strain can evolve to have both mutations and be resistant to both antibiotics. That is what happened in the case of Andrew Speaker.

He made his way to Denver for treatment, this time with a third antibiotic, kanamycin, which binds to the bacterial ribosome. The ribosome is the cell's protein synthesis factory, and is also essential for gene expression. Finally, the treatment was successful.

Mycobacterium tuberculosis The causative agent of TB can be killed with antibiotics, but resistance sometimes occurs.





Tuberculosis Is a World Health Issue Drug-resistant TB has become a major medical problem throughout the world. Here a doctor examines a patient in Ethiopia, which ranks high among the world's nations in number of TB cases.

Proteins are the major products of gene expression. Some proteins play vital structural roles in cells, and others act as enzymes, which are essential for most aspects of phenotypic expression. So, when protein synthesis is inhibited, cells cannot survive. This is what happened to the TB-causing bacteria when Speaker was treated with kanamycin. But genes can mutate, and the alleles that result may encode proteins that have altered surfaces. The mutant alleles in resistant TB encoded proteins that would no longer bind antibiotics.

IN THIS CHAPTER we will describe how genes are expressed as proteins, first discussing the evidence for the relationship between genes and proteins. We will then describe how the DNA sequence of a gene is copied (transcribed) into a sequence of RNA, and how the RNA sequence is translated to make a polypeptide with a defined sequence of amino acids. We will discuss some of the modifications to proteins that occur after they are made by the ribosomes. Following Mendel's definition of the gene as a physical entity, scientists characterized the gene as DNA (see Chapter 13). In this chapter we see how the gene gets expressed as a phenotype at the molecular level.

CHAPTER OUTLINE

- 14.1** What Is the Evidence that Genes Code for Proteins?
- 14.2** How Does Information Flow from Genes to Proteins?
- 14.3** How Is the Information Content in DNA Transcribed to Produce RNA?
- 14.4** How Is Eukaryotic DNA Transcribed and the RNA Processed?
- 14.5** How Is RNA Translated into Proteins?
- 14.6** What Happens to Polypeptides after Translation?

14.1 What Is the Evidence that Genes Code for Proteins?

In Chapter 12, we defined genes as sequences of DNA and learned that genes are expressed as physical characteristics known as the phenotype. Here, we show that in most cases, genes code for proteins, and it is the proteins that determine the phenotype. What is the evidence for this?

The molecular basis of phenotypes was actually discovered before it was known that DNA was the genetic material. Scientists had studied the chemical differences between individuals carrying wild-type and mutant alleles in organisms as diverse as humans and bread molds. They found that the major phenotypic differences resulted from differences in specific proteins.

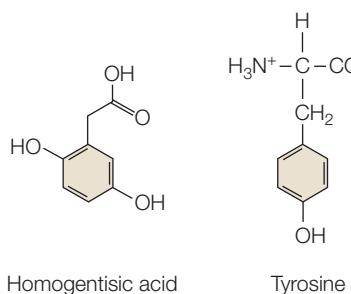
Observations in humans led to the proposal that genes determine enzymes

The identification of a gene product as a protein began with a mutation. In the early twentieth century, the English physician Archibald Garrod saw a number of children with a rare disease. One symptom of the disease was that the urine turned dark brown in air. This was especially noticeable on the infants' diapers. The disease was given the descriptive name alkapttonuria ("black urine").

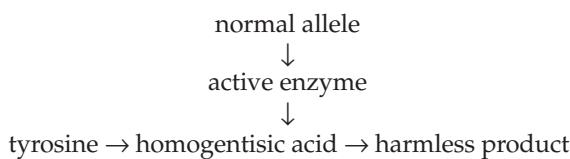
Garrod noticed that the disease was most common in children whose parents were first cousins. Mendelian genetics had just been "rediscovered," and Garrod realized that because first cousins share alleles (can you calculate what fraction?), the children of first cousins might inherit a rare mutant allele from both parents. He proposed that alkapttonuria was a phenotype caused by a recessive, mutant allele.

Garrod took the analysis one step further. He identified the biochemical abnormality in the affected children. He isolated from them an unusual substance, homogentisic acid, which accumulated in blood, joints (where it crystallized and caused severe pain), and urine (where it turned black). The chemical

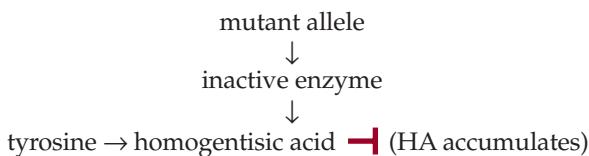
structure of homogentisic acid is similar to that of the amino acid tyrosine:



Enzymes as biological catalysts had just been discovered. Garrod proposed that homogentisic acid was a breakdown product of tyrosine. Normally, homogentisic acid is converted to a harmless product. According to Garrod, there is a normal (wild-type) human allele that determines the synthesis of an enzyme that catalyzes this conversion:



When the allele has been mutated, the enzyme is inactive and homogentisic acid accumulates instead:



Therefore, Garrod correlated *one gene to one enzyme* and coined the term “inborn error of metabolism” to describe this genetically determined biochemical disease. But his hypothesis needed direct confirmation by the identification of the specific enzyme and specific gene mutation involved. This did not occur until the enzyme, homogentisic acid oxidase, was described as active in healthy people and inactive in alkaptonuria patients in 1958, and the specific DNA mutation was described in 1996.

To directly relate genes and enzymes, biologists first turned to simpler organisms that could be manipulated in the laboratory.

Experiments on bread mold established that genes determine enzymes

As they work to explain the principles that underlie life, biologists often turn to organisms that they can manipulate experimentally. It wasn’t possible to follow up on Garrod’s hypothesis relating genes to enzymes in humans, because it is unethical to perform genetics experiments on people. Instead, biologists use *model organisms* that are easy to grow in the laboratory or greenhouse, and use them to develop principles of genetics that can then be applied more generally to other organisms. You have seen some of these model organisms in previous chapters:

- Pea plants (*Pisum sativum*) were used by Mendel in his genetics experiments.

- Fruit flies (*Drosophila*) were used by Morgan in his genetics experiments.
- *E. coli* was used by Meselson and Stahl to study DNA replication.

To this list we now add the common bread mold, *Neurospora crassa*. *Neurospora* is a type of fungus known as an ascomycete (see Chapter 30). This mold is haploid for most of its life, so that there are no dominant or recessive alleles: all alleles are expressed phenotypically and are not masked by a heterozygous condition. *Neurospora* is simple to culture and grows well in the laboratory. In the 1940s, George W. Beadle and Edward L. Tatum at Stanford University undertook studies to chemically define the phenotypes in *Neurospora*.

Beadle and Tatum knew about the roles of enzymes in biochemistry when they began their work, and like Garrod, they hypothesized that the expression of a specific gene results in the activity of a specific enzyme. Now, they set out to *test this hypothesis directly*. They grew *Neurospora* on a minimal nutritional medium containing sucrose, minerals, and a vitamin. Using this medium, the enzymes of wild-type *Neurospora* could catalyze all the metabolic reactions needed to make all the chemical constituents of their cells, including amino acids and proteins. These wild-type strains are called *prototrophs* (“original eaters”). From these wild type strains, they were able to produce and isolate distinct mutant strains that showed specific biochemical deficiencies.

Mutations provide a powerful way to determine cause and effect in biology. Nowhere has this been so evident as in the elucidation of biochemical pathways. Such pathways consist of sequential events (chemical reactions) in which each event is dependent on the occurrence of the preceding event. The general reasoning is as follows:

- *Observation.* Condition (1) occurs and condition (2) occurs; that is, (1) and (2) are *correlated*.
- *Hypothesis.* Condition (1) results in condition (2); that is, (1) *causes* (2).

In biochemical genetics, this can be stated as follows:

- *Observation.* A particular gene (*a*) is present and a particular reaction catalyzed by a particular enzyme (*A*) occurs; the two are correlated.
- *Hypothesis.* The gene (*a*) encodes (causes the synthesis of) the specific enzyme (*A*).
- *Test of hypothesis.* A mutant gene (*a'*) encodes a nonfunctional enzyme (*A'*) and the reaction does not occur.

Beadle and Tatum treated wild-type *Neurospora* with X rays, which act as a **mutagen** (something known to cause mutations— inherited genotypic changes). When they tested the treated molds, they found that some mutant strains could no longer grow on the minimal medium, but grew only if they were supplied with additional nutrients, such as amino acids. The scientists hypothesized that these *auxotrophs* (“increased eaters”) must have suffered mutations in genes that encoded the enzymes used to synthesize the nutrients that they now needed to obtain from their environment.

INVESTIGATING LIFE

14.1 One Gene, One Enzyme

Beadle and Tatum had several mutant strains of *Neurospora* that could not make arginine (*arg*). Several compounds are needed for arginine synthesis. By testing these compounds in the growth media for the mutant strains, the researchers deduced that each mutant strain was deficient in one enzyme along a biochemical pathway.

For each auxotrophic mutant strain, Beadle and Tatum were able to find a single compound that, when added to the minimal medium, supported the growth of that strain. These results suggested that mutations have simple effects, and that each mutation causes a defect in only one enzyme in a metabolic pathway. These conclusions confirmed Garrod's **one-gene, one-enzyme hypothesis** (Figure 14.1).

One group of auxotrophs, for example, could grow only if the minimal medium was supplemented with the amino acid arginine. These strains were designated *arg* mutants. Beadle and Tatum found several different *arg* mutant strains. They proposed two alternative hypotheses to explain why these different genetic strains had the same phenotype:

- The different *arg* mutants could have mutations in the same gene, as is the case for some eye color mutations in fruit flies. In this case, the gene might code for an enzyme involved in arginine synthesis.
- The different *arg* mutants could have mutations in different genes, each coding for a separate function that leads to arginine production. These independent functions might be different enzymes along the same biochemical pathway.

Some of the *arg* mutant strains fell into one of these two categories, and some into the other. Genetic crosses showed that some of the mutations were at the same chromosomal locus, and were different alleles of the same gene. Other mutations were at different loci, or on different chromosomes, and so were not alleles of the same gene. Beadle and Tatum concluded that these different genes participated in governing a single biosynthetic pathway—in this case, the pathway leading to arginine synthesis. Next, they set out to elucidate each step in this pathway (see the Interpretation in Figure 14.1).

By growing different *arg* mutants in the presence of various compounds suspected to be intermediates in the biosynthetic pathway for arginine, Beadle and Tatum were able to classify each mutation as affecting one enzyme or another, and to order the compounds along the pathway. Then they broke open the wild-type and mutant cells and examined them for enzyme activities. The results confirmed their hypothesis: each mutant strain was indeed missing a single enzyme activity in the pathway. In general, gene expression controls metabolism.

HYPOTHESIS Each gene determines an enzyme in a biochemical pathway.

METHOD

Place spores (single cells that divide to produce mold colonies) of each *arg* mutant strain on a minimal nutritional medium with and without supplements.

RESULTS

Supplement added to minimal medium

	None	Ornithine ↓	Citrulline ↓	Arginine ↓
Wild type				
Mutant strain 1				
Mutant strain 2				
Mutant strain 3				

The **wild type** grows on all media; it can synthesize its own arginine.

Mutant strain 1 grows only on arginine. It cannot convert either citrulline or ornithine to arginine.

Mutant strain 2 grows on either arginine or citrulline. It can convert citrulline to arginine, but cannot convert ornithine to citrulline.

Mutant strain 3 grows when any one of the three supplements are added. It can convert ornithine to citrulline and citrulline to arginine.

INTERPRETATION

Strain 3 is blocked here.

Gene A
↓
Enzyme A

Strain 2 is blocked here.

Gene B
↓
Enzyme B

Strain 1 is blocked here.

Gene C
↓
Enzyme C



If an organism cannot convert one particular compound to another, it presumably lacks an enzyme required for conversion, and the mutation is in the gene that codes for that enzyme.

CONCLUSION Each gene specifies a particular enzyme.

FURTHER INVESTIGATION: If a diploid *Neurospora* spore were made from two haploid cells, one with mutant 3 and the other with mutant 2, what would be its phenotype?

Go to yourBioPortal.com for original citations, discussions, and relevant links for all INVESTIGATING LIFE figures.

One gene determines one polypeptide

The gene–enzyme relationship has undergone several modifications in light of our current knowledge of molecular biology. Many proteins, including many enzymes, are composed of more than one polypeptide chain, or subunit (that is, they have a quaternary structure; see Section 3.2). Look at the illustration of hemoglobin in Figure 3.10. This protein has four polypeptides—two α and two β subunits, and the different subunits are encoded by separate genes. Thus it is more correct to speak of a **one-gene, one-polypeptide relationship**.

So far we have seen that in terms of protein synthesis, the *function of a gene is to inform the production of a single, specific polypeptide*. But not all genes code for polypeptides. As we will see below and in Chapter 16, there are many DNA sequences that code for RNA molecules that are not translated into polypeptides but instead have other functions.

14.1 RECAP

Beadle and Tatum's studies of mutations in bread molds led to our understanding of the one-gene, one-polypeptide relationship. In most cases, the function of a gene is to code for a specific polypeptide.

- What is a model organism, and why is *Neurospora* a good model for studying biochemical genetics?
See pp. 292
- How were Beadle and Tatum's experiments on *Neurospora* set up to determine the order of steps in a biochemical pathway? **See pp. 292–293 and Figure 14.1**
- Explain the distinction between the phrases "one-gene, one-enzyme" and "one-gene, one-polypeptide."
See pp. 293–294

Now that we have established the one-gene, one-polypeptide relationship, how does it work? That is, how is the information encoded in DNA used to produce a particular polypeptide?

14.2 How Does Information Flow from Genes to Proteins?

Much of the biochemical genetics in the middle of the twentieth century was directed at revealing the relationship between genes and protein synthesis. As we discussed in Section 14.1, the expression of a specific gene usually results in the synthesis of a specific polypeptide. The process of gene expression was outlined in Section 4.1. To review, this process occurs in two major steps:

- During **transcription**, the information in a DNA sequence (a gene) is copied into a complementary RNA sequence.
- During **translation**, this RNA sequence is used to create the amino acid sequence of a polypeptide.

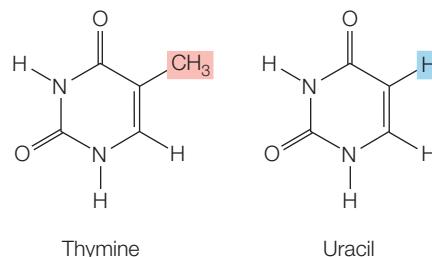


In 1958, Francis Crick described this process as "the **central dogma** of molecular biology."

RNA differs from DNA and plays a vital role in gene expression

RNA (ribonucleic acid) is a key intermediary between a DNA sequence and a polypeptide. RNA is an informational polynucleotide similar to DNA (see Figure 4.2), but it differs from DNA in three ways:

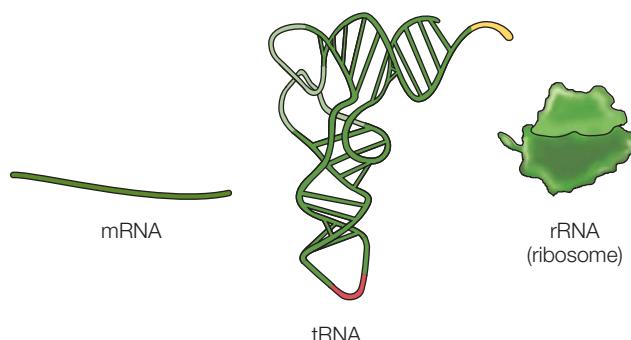
- RNA generally consists of only one polynucleotide strand.
- The sugar molecule found in RNA is ribose, rather than the deoxyribose found in DNA.
- Although three of the nitrogenous bases (adenine, guanine, and cytosine) in RNA are identical to those in DNA, the fourth base in RNA is **uracil (U)**, which is similar to thymine but lacks the methyl ($-\text{CH}_3$) group.

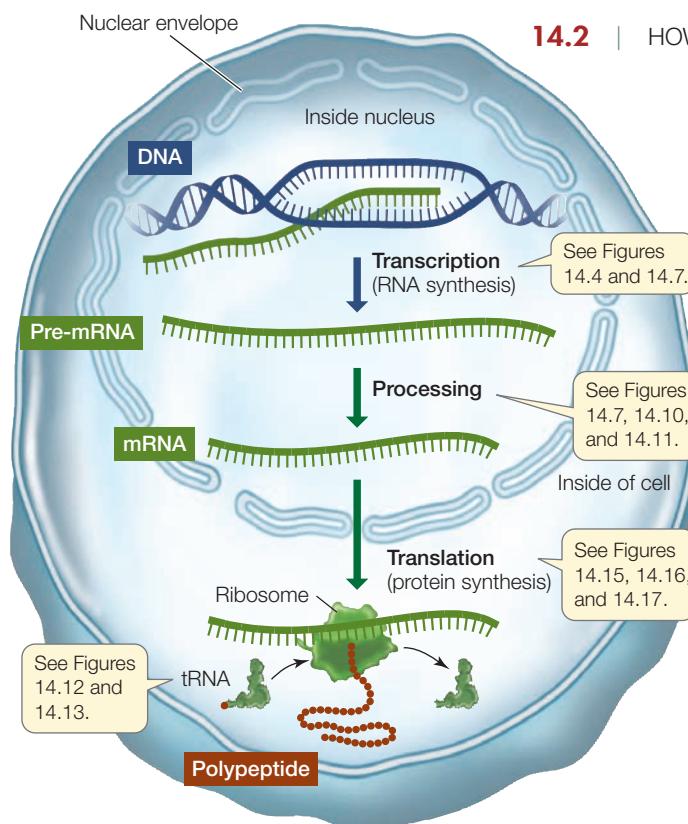


The bases in RNA can pair with those in a single strand of DNA. This pairing obeys the same complementary base-pairing rules as in DNA, except that adenine pairs with uracil instead of thymine. Single-stranded RNA can fold into complex shapes by internal base pairing, as seen below.

Three types of RNA participate in protein synthesis:

- **Messenger RNA (mRNA)** carries a copy of a gene sequence in DNA to the site of protein synthesis at the ribosome.
- **Transfer RNA (tRNA)** carries amino acids to the ribosome for assembly into polypeptides.
- **Ribosomal RNA (rRNA)** catalyzes peptide bond formation and provides a structural framework for the ribosome.





14.2 From Gene to Protein This diagram summarizes the processes of gene expression in eukaryotes.

yourBioPortal.com
GO TO Web Activity 14.1 • Eukaryotic Gene Expression

Two hypotheses were proposed to explain information flow from DNA to protein

The central dogma suggested that information flows from DNA to RNA to protein, and not in the reverse direction. It raised two questions:

- How does genetic information get from the nucleus of a eukaryotic cell to the cytoplasm? (As Section 5.3 explains, most of the DNA of a eukaryotic cell is confined to the nucleus, but proteins are synthesized in the cytoplasm.)
- What is the relationship between a specific nucleotide sequence in DNA and a specific amino acid sequence in a protein?

To answer these questions, Crick proposed two hypotheses, the messenger hypothesis and the adapter hypothesis.

THE MESSENGER HYPOTHESIS AND TRANSCRIPTION Crick and his colleagues proposed that an RNA molecule forms as a complementary copy of one DNA strand in a gene. This messenger RNA, or mRNA, then travels from the nucleus to the cytoplasm, where it serves as an informational sequence of **codons**. Each codon consists of three consecutive nucleotides, and different codons encode particular amino acids. Thus the mRNA sequence determines the ordered sequence of amino acids in a polypeptide chain, which is built by the ribosome. The process by which RNA forms is called transcription (Figure 14.2).

This hypothesis has been tested repeatedly, and the result is always the same: each DNA sequence that encodes a protein is transcribed as a sequence of mRNA. Today it is routine in thousands of laboratories around the world to test for gene expression by examining the mRNA copy of the gene, which is often called the **transcript**. There is no longer any question that Crick's model was correct.

THE ADAPTER HYPOTHESIS AND TRANSLATION To answer the question of how a DNA sequence gets transformed into the specific amino acid sequence of a polypeptide, Crick proposed the adapter hypothesis: that there must be an adapter molecule that can both bind a specific amino acid and recognize a specific sequence of nucleotides. He proposed that this recognition function occurs because the adapter molecule contains an **anticodon** complementary to the codon in the mRNA. He envisioned such adapters as molecules with two regions, one serving the binding function and the other serving the recognition function.

In due course, such adapter molecules were found: they are known as transfer RNA, or tRNA. Each tRNA recognizes a specific codon in the mRNA and simultaneously carries the specific amino acid corresponding to that codon. Thus, the tRNAs together can translate the language of DNA into the language of proteins. The tRNA adapters, carrying bound amino acids, line up on the mRNA sequence so that the amino acids are in the proper sequence for a growing polypeptide chain—in the process of *translation* (see Figure 14.2). Once again, actual observations of the expression of thousands of genes in all types of organisms have confirmed the hypothesis that tRNA acts as the intermediary between the nucleotide sequence information in mRNA and the amino acid sequence in a protein.

We can summarize the main features of the central dogma, the messenger hypothesis, and the adapter hypothesis as follows: a given gene is transcribed to produce an mRNA molecule that is complementary to one of the DNA strands, and then the tRNA molecules translate the sequence of codons in the mRNA into a sequence of linked amino acids, to form a polypeptide.

RNA viruses are exceptions to the central dogma

Certain viruses present exceptions to the central dogma. As we saw in Section 13.1, a virus is a non-cellular infectious particle that reproduces inside cells. Many viruses, such as the tobacco mosaic virus, influenza viruses, and poliovirus, have RNA rather than DNA as their genetic material. With its nucleotide sequence, RNA could potentially act as an information carrier and be expressed as a protein. But if RNA is usually single-stranded, how do these viruses replicate? They generally solve this problem by transcribing from RNA to RNA, making an RNA strand that is complementary to their genomes. This "opposite" strand is then used to make multiple copies of the viral genome by transcription:



Human immunodeficiency viruses (HIV) and certain rare tumor viruses also have RNA as their genomes, but do not replicate by transcribing from RNA to RNA. Instead, after infecting a host cell, such a virus makes a DNA copy of its genome, which becomes incorporated into the host's genome. The virus relies on the host cell's transcription machinery to make more RNA. This RNA can be either translated to produce viral proteins, or incorporated as the viral genome into new viral particles. Synthesis of DNA from RNA is called **reverse transcription**, and not surprisingly, such viruses are called **retroviruses**.

14.2 RECAP

The central dogma of molecular biology states that the DNA code is used to produce RNA and the RNA sequence determines the sequence of amino acids in a polypeptide. Transcription is the process by which a DNA sequence is copied into mRNA. Translation is the process by which this information is converted into polypeptide chain. Transfer RNAs recognize the genetic information in messenger RNA and bring the appropriate amino acids into position in a growing polypeptide chain.

- What is the central dogma of molecular biology? See p. 294
- What are the roles of mRNA and tRNA in gene expression? See p. 294 and Figure 14.2

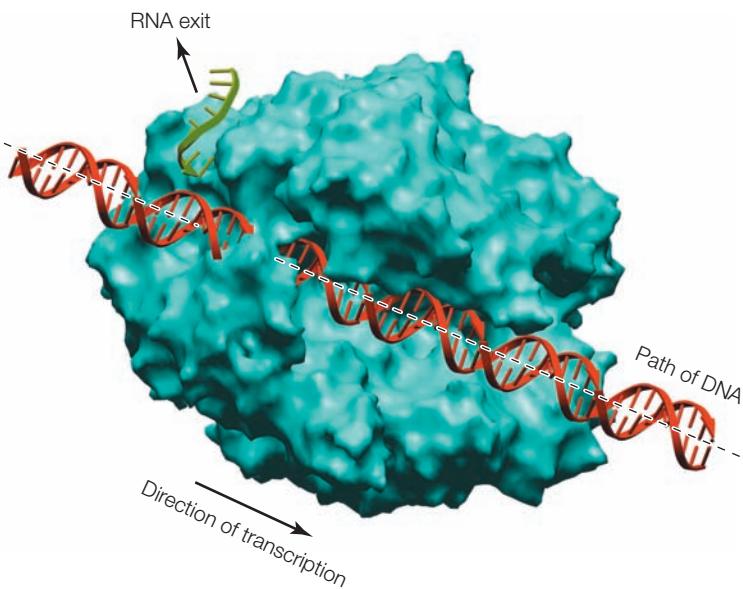
The central dogma is indeed central to gene expression in all organisms. Understanding its details is essential for understanding how organisms function at the molecular level, and this understanding is key to the application of biology to human welfare, in areas such as agriculture and medicine. Much of the remainder of this book will in one way or another involve DNA and proteins. Let's begin by describing how the information in DNA is transcribed to produce RNA.

14.3 How Is the Information Content in DNA Transcribed to Produce RNA?

In normal prokaryotic and eukaryotic cells, RNA synthesis is directed by DNA. Transcription—the formation of a specific RNA sequence from a specific DNA sequence—requires several components:

- A DNA template for complementary base pairing; one of the two strands of DNA
- The appropriate nucleoside triphosphates (ATP, GTP, CTP, and UTP) to act as substrates
- An RNA polymerase enzyme

Not only mRNA is produced by transcription. The same process is responsible for the synthesis of tRNA and ribosomal RNA (rRNA), whose important roles in protein synthesis will be described below. Like polypeptides, these RNAs are encoded by specific genes. In addition, as we will see in Chapter 16,



14.3 RNA Polymerase This enzyme from yeast is similar to most other RNA polymerases. Note the size relationship between enzyme and DNA. See Figure 14.4 for details.

many small RNAs, called microRNAs, are transcribed. These molecules stay in the nucleus, where they play roles in stimulating or inhibiting gene expression.

RNA polymerases share common features

RNA polymerases from both prokaryotes and eukaryotes catalyze the synthesis of RNA from the DNA template. There is only one kind of RNA polymerase in bacteria, while there are several kinds in eukaryotes; however, they all share a common structure (Figure 14.3). Like DNA polymerases, RNA polymerases are *processive*; that is, a single enzyme-template binding event results in the polymerization of hundreds of RNA bases. But unlike DNA polymerases, RNA polymerases *do not require a primer* and *do not have a proofreading function*.

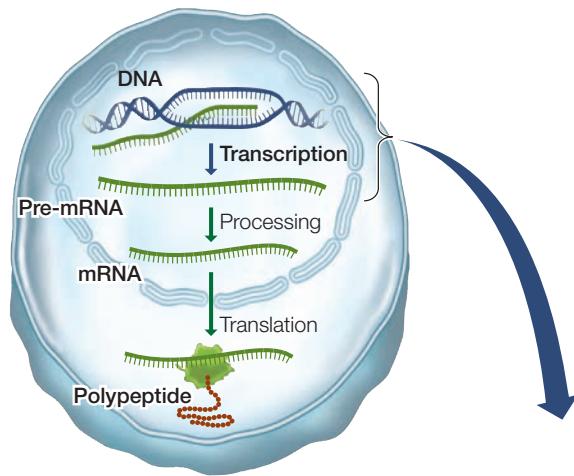
Transcription occurs in three steps

Transcription can be divided into three distinct processes: initiation, elongation, and termination. You can follow these processes in Figure 14.4.

INITIATION Transcription begins with initiation, which requires a **promoter**, a special sequence of DNA to which the RNA polymerase binds very tightly (see Figure 14.4A). Eukaryotic genes generally have one promoter each, while in prokaryotes and viruses, several genes often share one promoter. Promoters are important control sequences that “tell” the RNA polymerase two things:

- Where to start transcription
- Which strand of DNA to transcribe

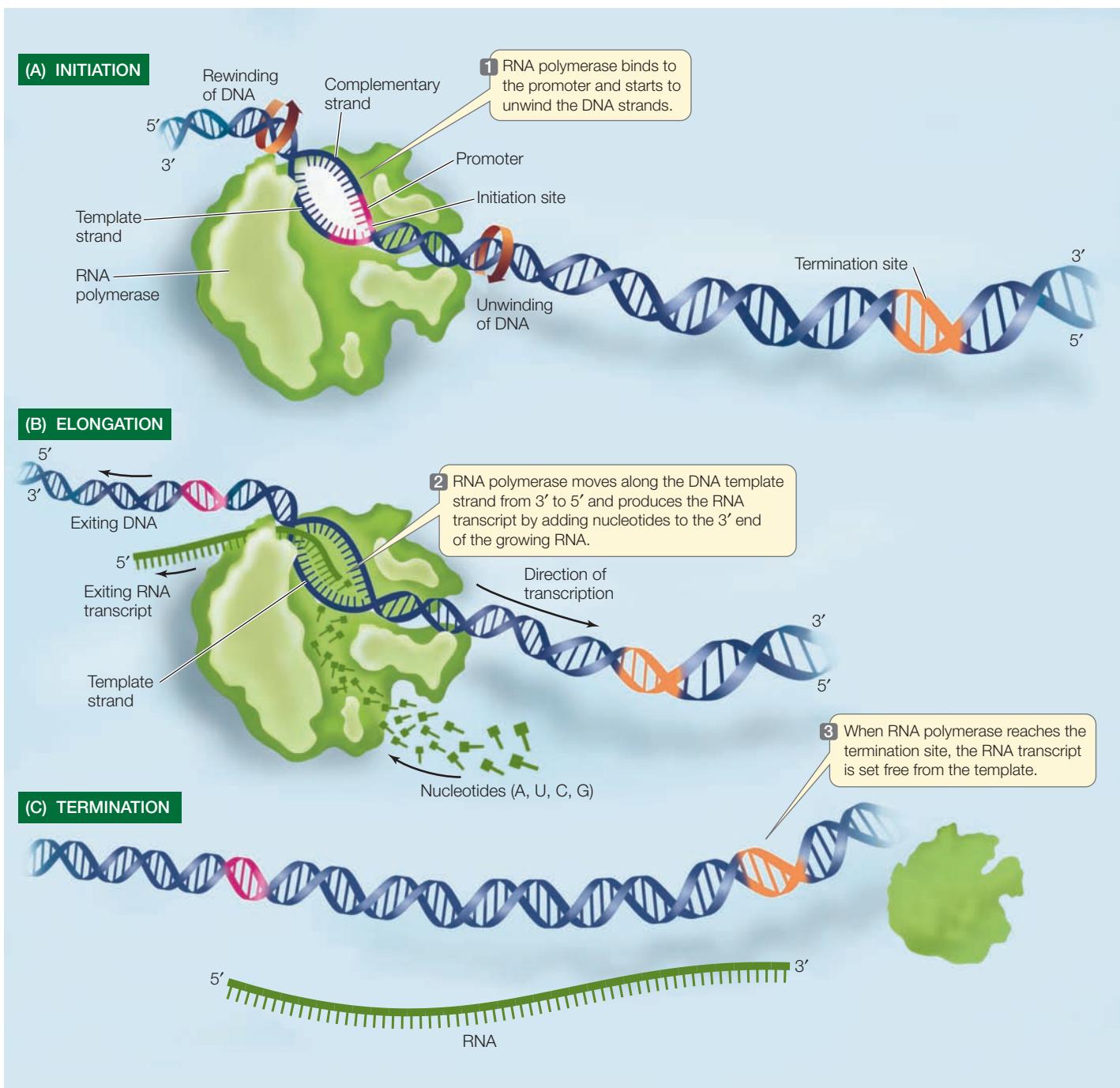
A promoter, which is a specific sequence in the DNA that reads in a particular direction, orients the RNA polymerase and thus “aims” it at the correct strand to use as a template. Promoters



14.4 DNA Is Transcribed to Form RNA DNA is partially unwound by RNA polymerase to serve as a template for RNA synthesis. The RNA transcript is formed and then peels away, allowing the DNA that has already been transcribed to rewind into a double helix. Three distinct processes—initiation, elongation, and termination—constitute DNA transcription. RNA polymerase is much larger in reality than indicated here, covering about 50 base pairs.

yourBioPortal.com

GO TO Animated Tutorial 14.1 • Transcription



function somewhat like the capital letter at the beginning of a sentence, indicating how the sequence of words should be read. Part of each promoter is the **initiation site**, where transcription begins. Groups of nucleotides lying “upstream” from the initiation site (5' on the non-template strand, and 3' on the template strand) help the RNA polymerase bind.

Although every gene has a promoter, not all promoters are identical. Some promoters are more effective at transcription initiation than others. Furthermore, there are differences between transcription initiation in prokaryotes and in eukaryotes. But despite these variations, the basic mechanisms of initiation are the same throughout the living world.

ELONGATION Once RNA polymerase has bound to the promoter, it begins the process of **elongation** (see Figure 14.4B). RNA polymerase unwinds the DNA about 10 base pairs at a time and reads the template strand in the 3'-to-5' direction. Like DNA polymerase, RNA polymerase adds new nucleotides to the 3' end of the growing strand, but does not require a primer to get this process started. The new RNA elongates from the first base, which forms its 5' end, to its 3' end. The RNA transcript is thus antiparallel to the DNA template strand.

Because RNA polymerases do not proofread, transcription errors occur at a rate of one for every 10^4 to 10^5 bases. Because many copies of RNA are made, however, and because they often have only a relatively short life span, these errors are not as potentially harmful as mutations in DNA.

TERMINATION Just as initiation sites in the DNA template strand specify the starting point for transcription, particular base sequences specify its **termination** (see Figure 14.4C). The mechanisms of termination are complex and of more than one kind. For some genes, the newly formed transcript falls away from the DNA template and the RNA polymerase. For others, a helper protein pulls the transcript away.

The information for protein synthesis lies in the genetic code

The **genetic code** relates genes (DNA) to mRNA and mRNA to the amino acids that make up proteins. The genetic code specifies which amino acids will be used to build a protein. You can think of the genetic information in an mRNA molecule as a series of sequential, nonoverlapping three-letter “words.” Each sequence of three nucleotide bases (the three “letters”) along the mRNA polynucleotide chain specifies a particular amino acid. Each three-letter “word” is called a **codon**. Each

codon is complementary to the corresponding triplet of bases in the DNA molecule from which it was transcribed. The genetic code relates codons to their specific amino acids.

CHARACTERISTICS OF THE CODE Molecular biologists “broke” the genetic code in the early 1960s. The problem they addressed was perplexing: how could more than 20 “code words” be written with an “alphabet” consisting of only four “letters”? In other words, how could four bases (A, U, G, and C) code for 20 different amino acids?

A triplet code, based on three-letter codons, was considered likely. Since there are only four letters (A, G, C, and U), a one-letter code clearly could not unambiguously encode 20 amino acids; it could encode only four of them. A two-letter code could have only $4 \times 4 = 16$ unambiguous codons—still not enough. But a triplet code could have $4 \times 4 \times 4 = 64$ codons, more than enough to encode the 20 amino acids.

Marshall W. Nirenberg and J. H. Matthaei, at the U.S. National Institutes of Health, made the first decoding breakthrough in 1961 when they realized that they could use a simple artificial polynucleotide instead of a complex natural mRNA as a messenger. They could then identify the polypeptide that the artificial messenger encoded. This led to the identification of the first three codons (Figure 14.5).

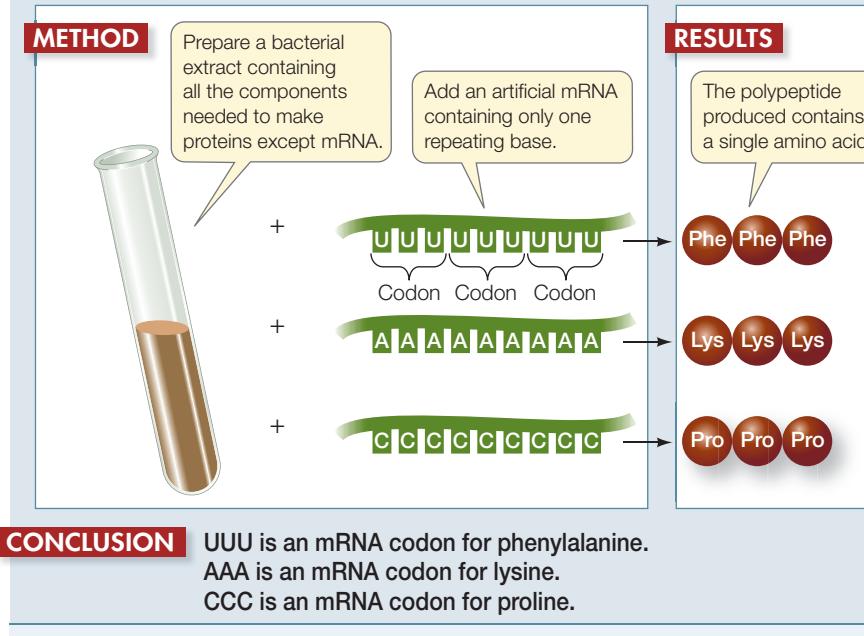
INVESTIGATING LIFE

14.5 Deciphering the Genetic Code

Nirenberg and Matthaei used a test-tube protein synthesis system to determine the amino acids specified by synthetic mRNAs of known codon compositions.

HYPOTHESIS

A triplet codon based on three-base codons specifies amino acids.



FURTHER INVESTIGATION: What would be the result if the artificial mRNA were poly-G?

Go to yourBioPortal.com for original citations, discussions, and relevant links for all INVESTIGATING LIFE figures.

Other scientists later found that simple artificial mRNAs only three nucleotides long—each amounting to one codon—could bind to a ribosome, and that the resulting complex could then bind to the corresponding tRNA with its specific amino acid. Thus, for example, a simple UUU mRNA caused the tRNA carrying phenylalanine to bind to the ribosome. After this discovery, the complete deciphering of the genetic code was relatively simple. To discover which amino acid a codon represented, the scientists simply repeated the experiment using a sample of artificial mRNA for that codon, and observed which amino acid became bound to it.

The complete genetic code is shown in **Figure 14.6**. Notice that there are many more codons than there are different amino acids in proteins. All possible combinations of the four available “letters” (the bases) give 64 (4^3) different three-letter codons, yet these codons determine only 20 amino acids. AUG, which codes for methionine, is also the **start codon**, the initiation signal for translation. Three of the codons (UAA, UAG, UGA) are **stop codons**, or termination signals for translation. When the translation machinery reaches one of these codons, translation stops, and the polypeptide is released from the translation complex.

What happens if a stop codon isn’t there? In humans a severe anemic condition, α -thalassemia, results from various mutations in either of the two genes that encode the α -polypeptide chain of hemoglobin. In one of these mutant alleles, the stop codon UAA has been converted to GAA, a codon for glutamine. The next stop codon doesn’t occur until much further along the mRNA, resulting in a protein molecule with larger, defective α subunits.

THE GENETIC CODE IS REDUNDANT BUT NOT AMBIGUOUS The 60 codons that are not start or stop codons are far more than enough to code for the other 19 amino acids—and indeed, for almost all amino acids, there is more than one codon. Thus we say that the genetic code is redundant (or degenerate). For ex-

ample, leucine is represented by six different codons (see Figure 14.6). Only methionine and tryptophan are represented by just one codon each.

A *redundant* code should not be confused with an *ambiguous* code. If the code were ambiguous, a single codon could specify either of two (or more) different amino acids, and there would be doubt about which amino acid should be incorporated into a growing polypeptide chain. Redundancy in the code simply means that there is more than one clear way to say, “Put leucine here.” The genetic code is not ambiguous: a given amino acid may be encoded by more than one codon, but a codon can code for only one amino acid.

THE GENETIC CODE IS (NEARLY) UNIVERSAL The same genetic code is used by all the species on our planet. Thus the code must be an ancient one that has been maintained intact throughout the evolution of living organisms. Exceptions are known: within mitochondria and chloroplasts, the code differs slightly from that in prokaryotes and in the nuclei of eukaryotic cells; and in one group of protists, UAA and UAG code for glutamine rather than functioning as stop codons. The significance of these differences is not yet clear. What is clear is that the exceptions are few.

The common genetic code means that there is also a common language for evolution. Natural selection acts on phenotypic variations that result from genetic variation. The genetic code probably originated early in the evolution of life. As we saw in Chapter 4, simulation experiments indicate the plausibility of individual nucleotides and nucleotide polymers arising spontaneously on the primeval Earth. The common code also has profound implications for genetic engineering, as we will see in Chapter 18, since it means that the code for a human gene is the same as that for a bacterial gene. It is therefore impressive, but not surprising, that a human gene can be expressed in *E. coli* via laboratory manipulations, since these cells speak the same “molecular language.”

The codons in Figure 14.6 are mRNA codons. The base sequence of the DNA strand that is transcribed to produce the mRNA is complementary and antiparallel to these codons. Thus, for example,

- 3'-AAA-5' in the template DNA strand corresponds to phenylalanine (which is encoded by the mRNA codon 5'-UUU-3')

14.6 The Genetic Code Genetic information is encoded in mRNA in three-letter units—codons—made up of nucleoside monophosphates with the bases uracil (U), cytosine (C), adenine (A), and guanine (G) and is read in a 5' to 3' direction on mRNA. To decode a codon, find its first letter in the left column, then read across the top to its second letter, then read down the right column to its third letter. The amino acid the codon specifies is given in the corresponding row. For example, AUG codes for methionine, and GUA codes for valine.

		Second letter					
		U	C	A	G		
First letter	U	UUU UUC UUA UUG	UCU UCC UCA UCG	UAU UAC	UGU UGC	Cysteine	U C
	C	CUU CUC CUA CUG	CCU CCC CCA CCG	CAU CAC	CGU CGC CGA CGG	Arginine	A G
A	C	CUU CUC CUA CUG	CCU CCC CCA CCG	CAU CAC	CGU CGC CGA CGG	Arginine	U C
	A	AUU AUC AUA AUG	ACU ACC ACA ACG	AAU AAC	AGU AGC	Serine	A G
G	A	AUU AUC AUA AUG	ACU ACC ACA ACG	AAU AAC	AGU AGC	Serine	U C
	G	GUU GUC GUA GUG	GCU GCC GCA GCG	AAA AAG	AGA AGG	Arginine	A G
		Third letter					

- 3'-ACC-5' in the template DNA corresponds to tryptophan (which is encoded by the mRNA codon 5'-UGG-3')

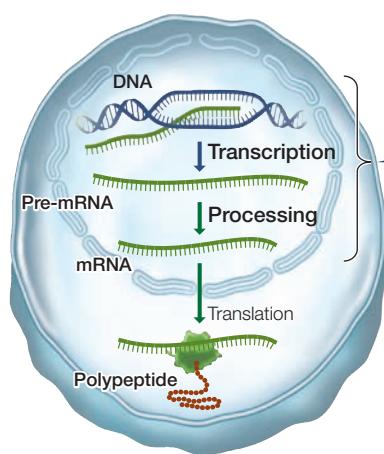
The non-template strand has the same sequence as the mRNA (but with T's instead of U's), and is often referred to as the "coding strand." By convention, DNA sequences are usually shown beginning with the 5' end of the coding sequence.

14.3 RECAP

Transcription, which is catalyzed by an RNA polymerase, proceeds in three steps: initiation, elongation, and termination. The genetic code relates the information in mRNA (as a linear sequence of codons) to protein (a linear sequence of amino acids).

- What are the steps of gene transcription that produce mRNA? See pp. 296–298 and Figure 14.4
- How do RNA polymerases work? See pp. 296–298
- How was the genetic code elucidated? See pp. 298–299 and Figure 14.5

The features of transcription that we have described were first elucidated in model prokaryotes, such as *E. coli*. Biologists then used the same methods to analyze this process in eukaryotes, and, although the basics are the same, there are some notable (and important) differences. We now turn to eukaryotic gene expression.



14.7 Transcription of a Eukaryotic Gene The β -globin gene diagrammed here is about 1,600 base pairs (bp) long. The three exons—the protein-coding sequences—contain codons for 146 amino acids plus a stop codon. The two introns—noncoding sequences of DNA containing almost 1,000 bp between them—are initially transcribed, but are spliced out of the pre-mRNA transcript.

14.4 How Is Eukaryotic DNA Transcribed and the RNA Processed?

Since the genetic code is the same, you might expect the process of gene expression to be the same in eukaryotes as it is in prokaryotes. And basically it is. However, there are significant differences in gene structure between prokaryotes and eukaryotes, that is, there are differences in the organization of the nucleotide sequences in the genes. In addition, in eukaryotes but not prokaryotes, a nucleus separates transcription and translation (Table 14.1). Let's look at the distinctive *eukaryotic* process of transcription.

Eukaryotic genes have noncoding sequences

A diagram of the structure and transcription of a typical eukaryotic gene is shown in Figure 14.7. In prokaryotes, several adjacent genes sometimes share one promoter; however, in eukaryotes, each gene has its own promoter, which usually precedes the coding region. Unlike the prokaryotic RNA polymerase, a eukaryotic RNA polymerase does not recognize the promoter sequence by itself, but requires help from other molecules, as we'll see in more detail in Chapter 16. At the other end of the gene, downstream from the coding region, is a DNA sequence appropriately called the **terminator**, which signals the end of transcription.

Eukaryotic genes may also contain noncoding base sequences, called **introns** (*intervening regions*). One or more introns may be interspersed with the coding sequences, which are called **exons** (*expressed regions*). Both introns and exons appear in the primary mRNA transcript, called **pre-mRNA**, but the introns are removed by the time the mature mRNA—the mRNA that will be translated—leaves the nucleus. Pre-mRNA processing involves cutting introns out of the pre-mRNA transcript and splicing together the remaining exon transcripts (see Figure 14.7). If this seems surprising, you are in good company. For scientists who were familiar with prokaryotic genes and gene expression, the discovery of introns in eukaryotic genes was entirely unexpected.

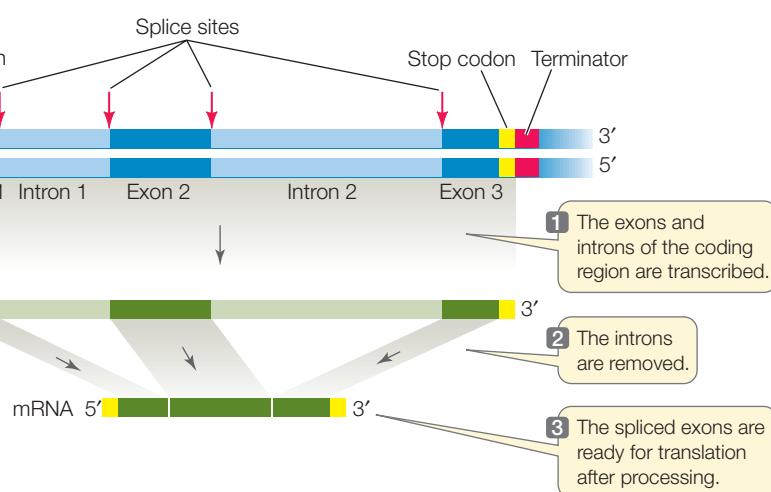


TABLE 14.1
Differences between Prokaryotic and Eukaryotic Gene Expression

CHARACTERISTIC	PROKARYOTES	EUKARYOTES
Transcription and translation occurrence	At the same time in the cytoplasm	Transcription in the nucleus, then translation in the cytoplasm
Gene structure	DNA sequence is read in the same order as the amino acid sequence	Noncoding introns within coding sequence
Modification of mRNA after initial transcription but before translation	None	Introns spliced out; 5' cap and 3' poly A added

How can we locate introns within a eukaryotic gene? One way is by **nucleic acid hybridization**, the method that originally revealed the existence of introns. This method, outlined in **Figure 14.8**, has been crucial for studying the relationship between eukaryotic genes and their transcripts. It involves two steps:

- The target DNA is denatured by heat to break the hydrogen bonds between the base pairs and separate the two strands.
- A single-stranded nucleic acid from another source (called a **probe**) is incubated with the denatured DNA. If the probe has a base sequence complementary to the target DNA, a probe–target double helix forms by hydrogen bonding between the bases. Because the two strands are from different sources, the resulting double-stranded molecule is called a hybrid.

Biologists used nucleic acid hybridization to examine the β -globin gene, which encodes one of the globin polypeptides that make up hemoglobin. Follow the experiment in **Figure 14.9** carefully as we describe what they did and what happened.

The researchers first denatured DNA containing the β -globin gene by heating it slowly, then added previously isolated, mature β -globin mRNA. They were able to view the hybridized molecules using electron microscopy. As expected, the mRNA bound to the DNA by complementary base pairing. The researchers expected to obtain a linear (1:1) matchup of the mRNA to the coding DNA. That expectation was only partly met: there were indeed stretches of RNA–DNA hybrid, but some looped structures were also visible. These loops were not expected, and initially the scientists thought that something must be wrong with the experimental procedure. However, when they repeated the experiment, they got the same results, and when they did it with other genes and mRNAs, loops again appeared. The loops turned out to be the introns, stretches of DNA that did not have complementary base sequences on the mature mRNA.

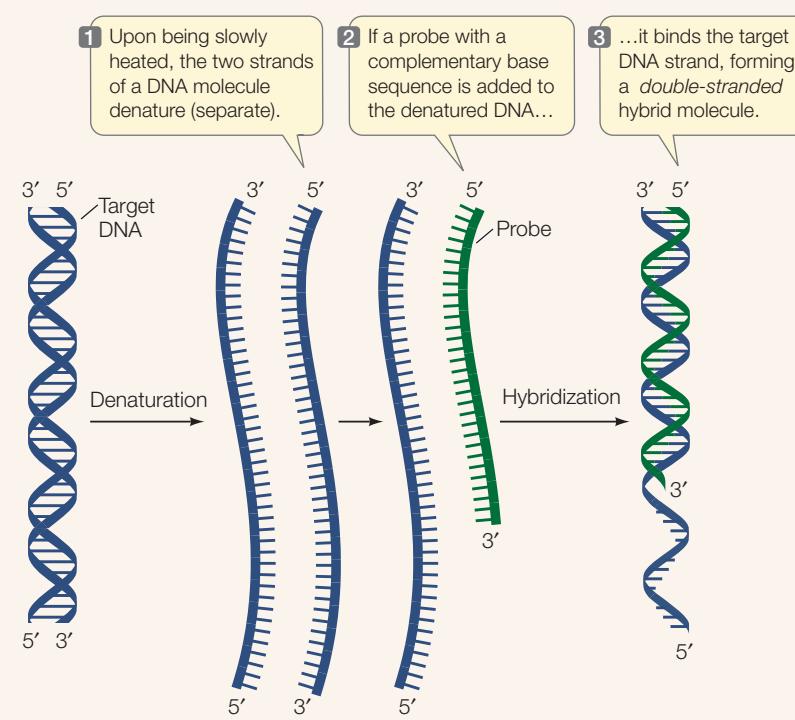
When pre-mRNA was used instead of mature mRNA to hybridize to the DNA, there was complete hybridization, revealing that the introns were indeed part of the pre-mRNA transcript. Somewhere on the path from primary transcript (pre-mRNA) to mature mRNA, the introns had been removed, and the exons had been spliced together. We will examine this splicing process in the next section.

Introns *interrupt, but do not scramble*, the DNA sequence of a gene. The base sequences of the exons in the template strand, if joined and taken in order, form a continuous sequence that is complementary to that of the mature mRNA. In some cases, the separated exons encode different functional regions, or **domains**, of the protein. For example, the globin polypeptides that make up hemoglobin each have two domains: one for binding to a nonprotein pigment called heme, and another for binding to the other globin subunits. These two domains are encoded by different exons in the globin genes. Most (but not all) eukaryotic genes contain introns, and in rare cases, introns are also found in prokaryotes. The largest human gene encodes a muscle protein called titin; it has 363 exons, which together code for 38,138 amino acids.

TOOLS FOR INVESTIGATING LIFE

14.8 Nucleic Acid Hybridization

Base pairing permits the detection of a sequence that is complementary to the probe.

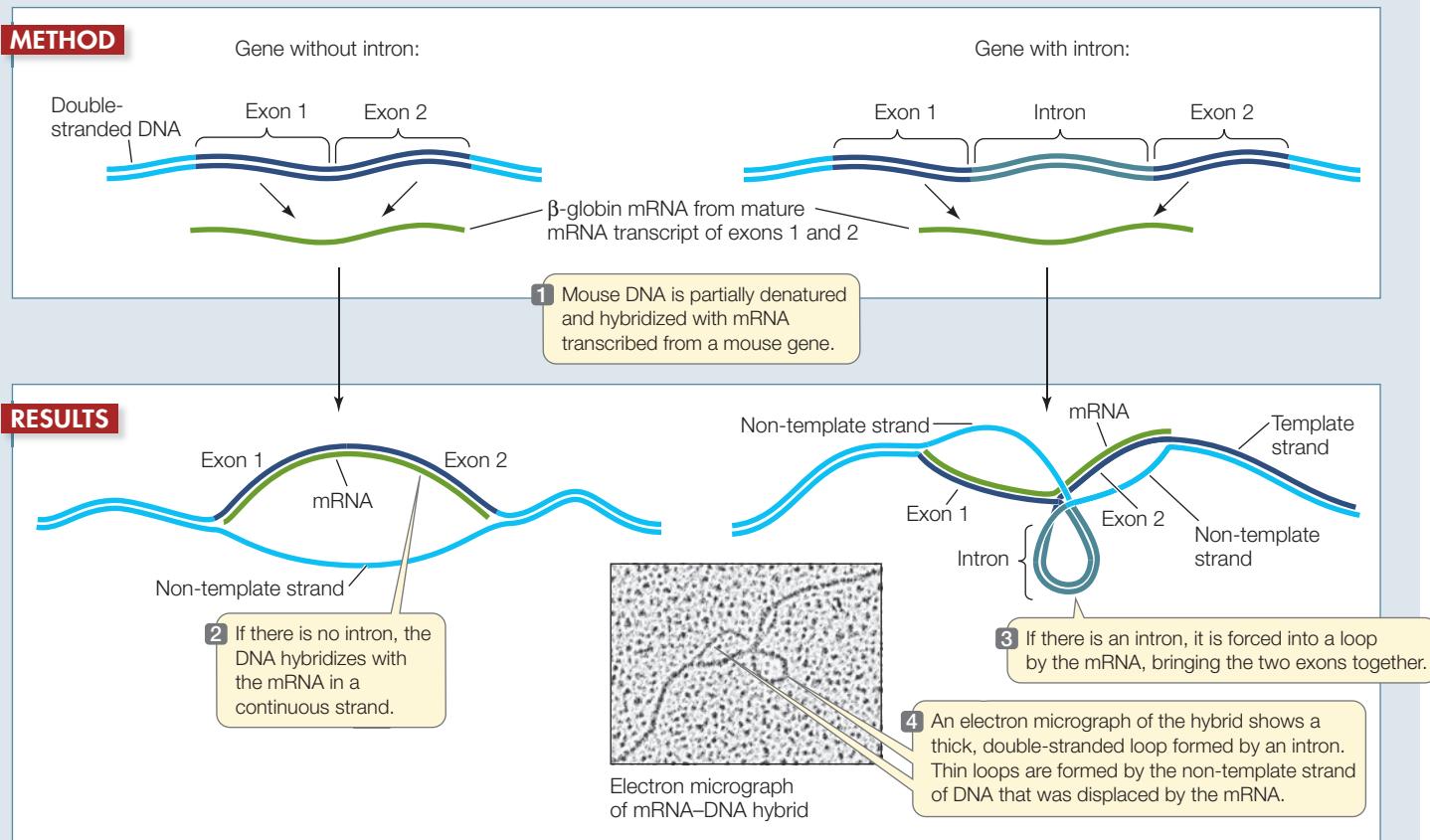


INVESTIGATING LIFE

14.9 Demonstrating the Existence of Introns

When an mRNA transcript of the β -globin gene was hybridized with the double-stranded DNA of that gene, the introns in the DNA “looped out.” This demonstrated that the coding region of a eukaryotic gene can contain noncoding DNA that is not present in the mature mRNA transcript.

HYPOTHESIS Some regions within the coding sequence of a gene do not end up in its mRNA.



CONCLUSION The DNA contains noncoding regions within the genes that are not present in the mature mRNA.

FURTHER INVESTIGATION: Draw the result assuming that there were three exons and two introns.

Go to yourBioPortal.com for original citations, discussions, and relevant links for all INVESTIGATING LIFE figures.

Eukaryotic gene transcripts are processed before translation

The primary transcript of a eukaryotic gene is modified in several ways before it leaves the nucleus: both ends of the pre-mRNA are modified, and the introns are removed.

MODIFICATION AT BOTH ENDS Two steps in the processing of pre-mRNA take place in the nucleus, one at each end of the molecule (**Figure 14.10**):

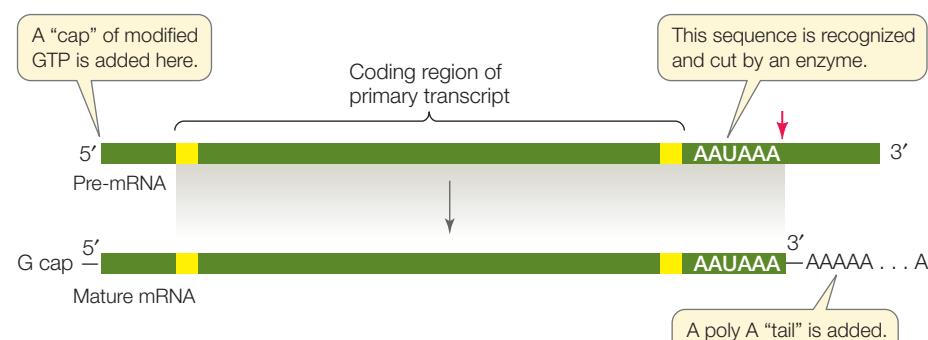
- A **G cap** is added to the 5' end of the pre-mRNA as it is transcribed. The G cap is a chemically modified molecule of guanosine triphosphate (GTP). It facilitates the binding of

mRNA to the ribosome for translation, and it protects the mRNA from being digested by ribonucleases that break down RNAs.

- A **poly A tail** is added to the 3' end of the pre-mRNA at the end of transcription. In both prokaryotic and eukaryotic genes, transcription begins at a DNA sequence that is upstream (to the “left” on the DNA) of the first codon (i.e., at the promoter), and ends downstream (to the “right” on the DNA) of the termination codon. In eukaryotes, there is usually a “polyadenylation” sequence (AAUAAA) near the 3' end of the pre-mRNA, after the last codon. This sequence acts as a signal for an enzyme to cut the pre-mRNA. Immediately after this cleavage, another enzyme

14.10 Processing the Ends of Eukaryotic Pre-mRNA

Pre-mRNA Modifications at each end of the pre-mRNA transcript—the G cap and the poly A tail—are important for mRNA function.



adds 100 to 300 adenine nucleotides (a “poly A” sequence) to the 3’ end of the pre-mRNA. This “tail” may assist in the export of the mRNA from the nucleus and is important for mRNA stability.

SPLICING TO REMOVE INTRONS The next step in the processing of eukaryotic pre-mRNA within the nucleus is removal of the introns. If these RNA sequences were not removed, a very different amino acid sequence, and possibly a nonfunctional protein, would result. A process called **RNA splicing** removes the introns and splices the exons together.

As soon as the pre-mRNA is transcribed, several **small nuclear ribonucleoprotein particles (snRNPs)** bind at each end. There are several types of these RNA–protein particles in the nucleus.

At the boundaries between introns and exons are **consensus sequences**—short stretches of DNA that appear, with little variation (“consensus”), in many different genes. The RNA in one of the snRNPs has a stretch of bases complementary to the consensus sequence at the 5’ exon–intron boundary, and it binds to the pre-mRNA by complementary base pairing. Another snRNP binds to the pre-mRNA near the 3’ intron–exon boundary (Figure 14.11).

Next, using energy from adenosine triphosphate (ATP), proteins are added to form a large RNA–protein complex called a **spliceosome**. This complex cuts the pre-mRNA, releases the introns, and joins the ends of the exons together to produce mature mRNA.

Molecular studies of human genetic diseases have provided insights into intron consensus sequences and splicing machinery. For example, people with the genetic disease β -thalassemia, like those with α -thalassemia discussed earlier in the chapter, have a defect in the production of one of the

yourBioPortal.com

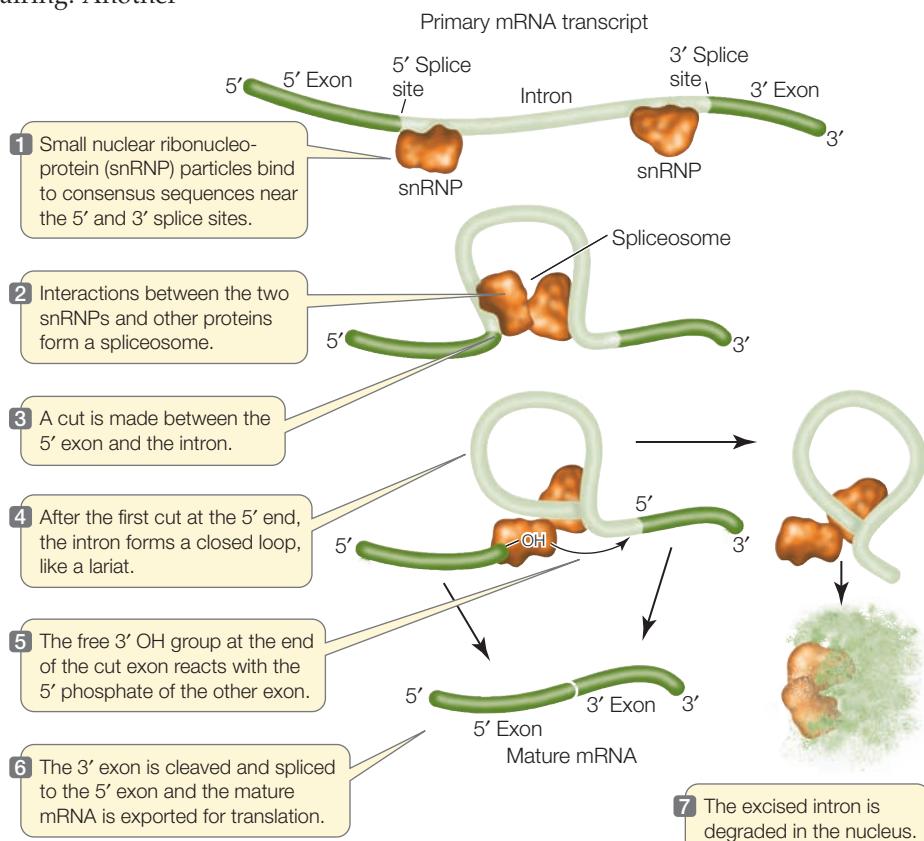
GO TO Animated Tutorial 14.3 • RNA Splicing

14.11 The Spliceosome: An RNA Splicing Machine

Machine The binding of snRNPs to consensus sequences bordering the introns on the pre-mRNA results in a series of proteins binding and forming a large complex called a spliceosome. This structure determines the exact position of each cut in the pre-mRNA with great precision.

hemoglobin subunits. These people suffer from severe anemia because they have an inadequate supply of red blood cells. In some cases, the genetic mutation that causes the disease occurs at an intron consensus sequence in the β -globin gene. Consequently, β -globin pre-mRNA cannot be spliced correctly, and nonfunctional β -globin mRNA is made. This finding offers another example of how biologists can use mutations to elucidate cause-and-effect relationships.

After processing is completed in the nucleus, the mature mRNA moves out into the cytoplasm through the nuclear pores. In the nucleus, a protein called TAP binds to the 5’ end of processed mRNA. This protein in turn binds to others, which are recognized by a receptor at the nuclear pore. Together, these proteins lead the mRNA through the pore. Unprocessed or incompletely processed pre-mRNAs remain in the nucleus.



14.4 RECAP

Most eukaryotic genes contain noncoding sequences called introns, which are removed from the pre-mRNA transcript.

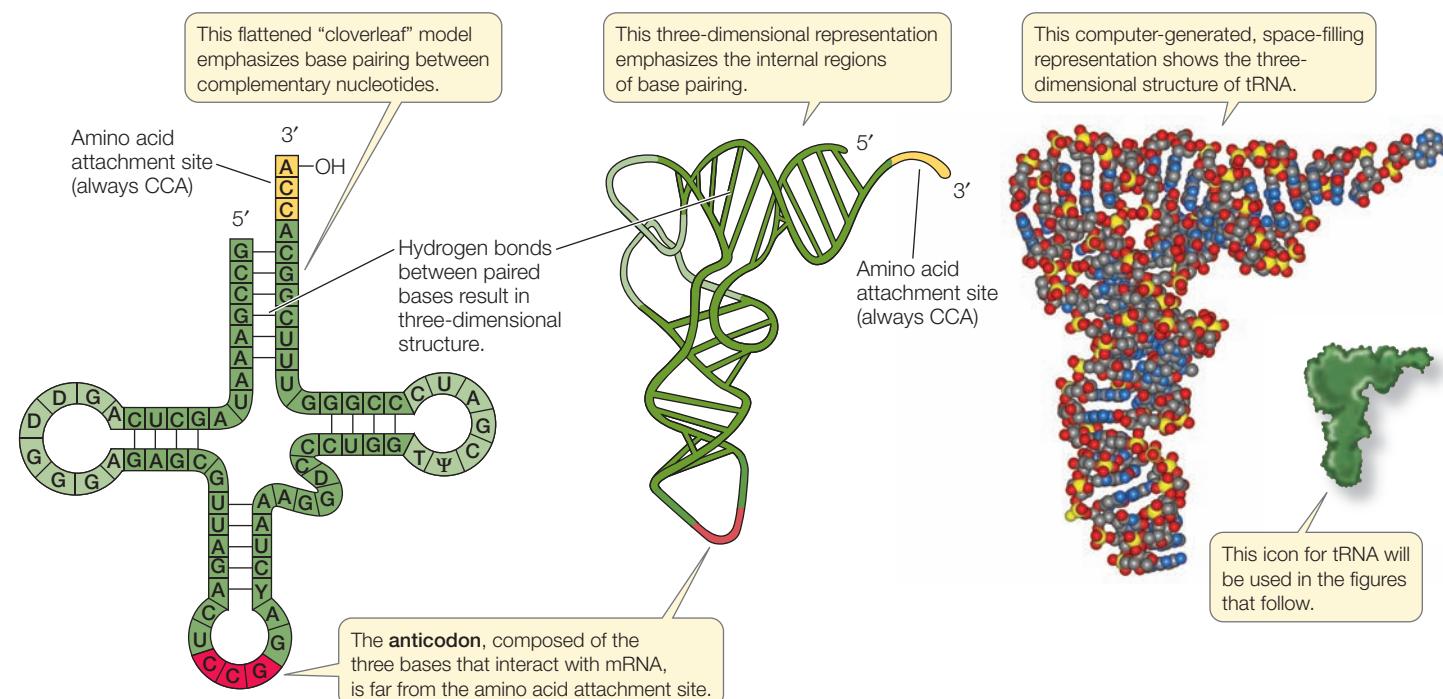
- Describe the method of nucleic acid hybridization. **See p. 301 and Figure 14.8**
- Describe the experiment that showed that the β -globin gene contains introns. **See p. 301 and Figure 14.9**
- How is the pre-mRNA transcript modified at the 5' and 3' ends? **See p. 302 and Figure 14.10**
- How does RNA splicing happen? What are the consequences if it does not happen correctly? **See p. 303 and Figure 14.11**

Transcription and post-transcriptional events produce an mRNA that is ready to be translated into a sequence of amino acids in a polypeptide. We turn now to the events of translation.

14.5 How Is RNA Translated into Proteins?

As Crick's adapter hypothesis proposed, the translation of mRNA into proteins requires a molecule that links the information contained in mRNA codons with specific amino acids in proteins. That function is performed by transfer RNA (tRNA). Two key events must take place to ensure that the protein made is the one specified by the mRNA:

- The tRNAs must read mRNA codons correctly.
- The tRNAs must deliver the amino acids that correspond to each mRNA codon.



Once the tRNAs "decode" the mRNA and deliver the appropriate amino acids, components of the ribosome catalyze the formation of peptide bonds between amino acids. We now turn to these two steps.

yourBioPortal.com

GO TO Animated Tutorial 14.4 • Protein Synthesis

Transfer RNAs carry specific amino acids and bind to specific codons

A codon in mRNA and an amino acid in a protein are related by way of an adapter—a specific tRNA with an attached amino acid. For each of the 20 amino acids, there is at least one specific type ("species") of tRNA molecule. The tRNA molecule has three functions:

- It binds to a particular amino acid. When it is carrying an amino acid, the tRNA is said to be "charged."
- It associates with mRNA.
- It interacts with ribosomes.

The tRNA molecular structure relates clearly to all of these functions. The molecule has about 75 to 80 nucleotides. It has a conformation (a three-dimensional shape) that is maintained by complementary base pairing (hydrogen bonding) between bases within its own sequence (**Figure 14.12**).

The conformation of a tRNA molecule is exquisitely suited for its interaction with specific binding sites on ribosomes. In addition, at the 3' end of every tRNA molecule is its amino acid attachment site: a site to which its specific amino acid binds co-

valently. At about the midpoint of the tRNA sequence is a group of three bases, called the anticodon, which is the site of complementary base pairing (via hydrogen bonding) with the codon on the mRNA. Thus, each tRNA species has a unique anticodon that corresponds to the amino acid it carries. When the tRNA and the mRNA come into contact on the surface of the ribosome, the codon and anticodon are antiparallel, permitting hydrogen bonding to occur between the complementary bases. As an example of this process, consider the amino acid arginine:

- The template strand DNA sequence that codes for arginine is 3'-GCC-5', which is transcribed, by complementary base pairing, to produce the mRNA codon 5'-CGG-3'
- That mRNA codon binds by complementary base pairing to a tRNA with the anticodon 3'-GCC-5', which is charged with arginine.

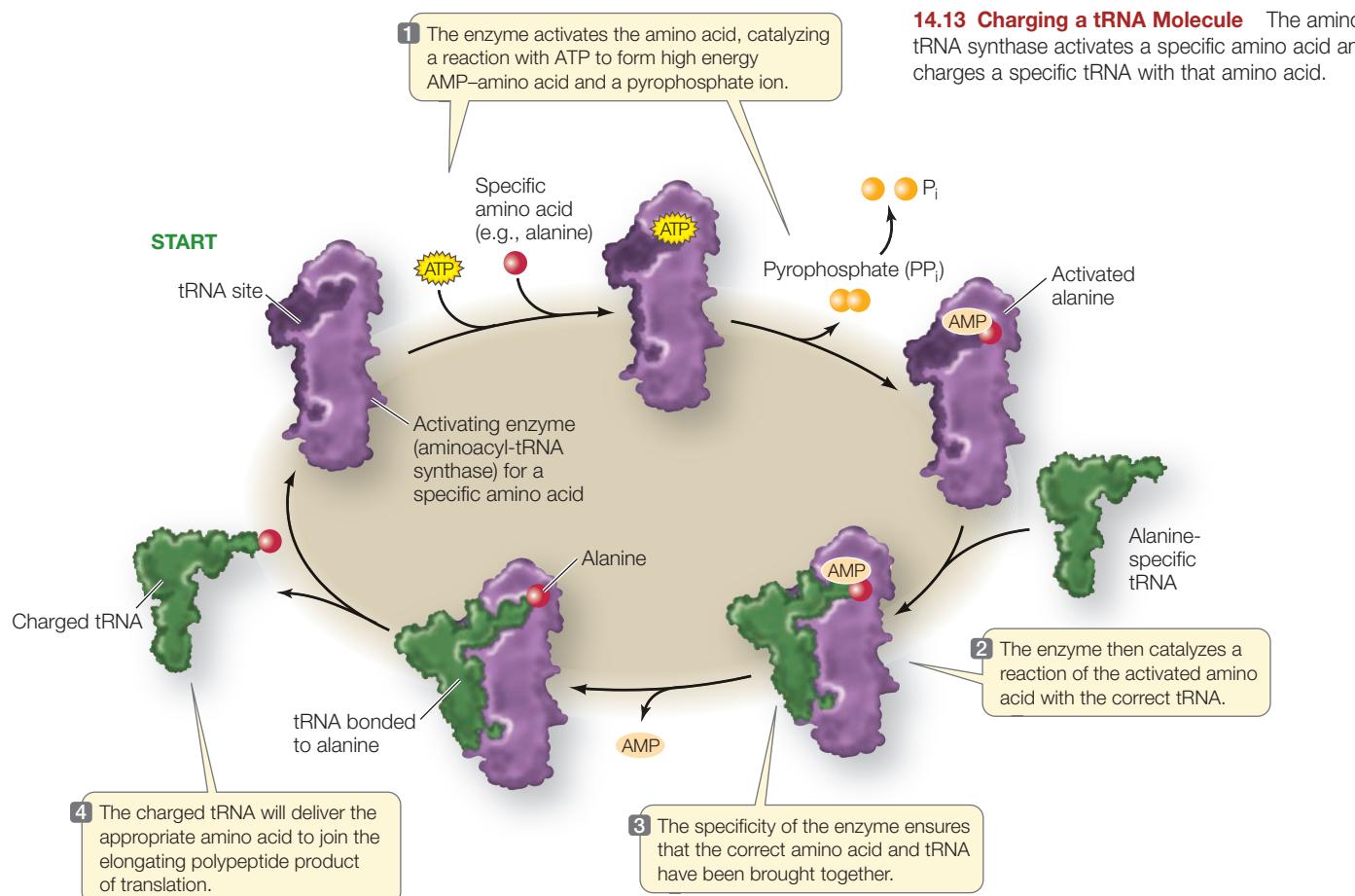
Recall that 61 different codons encode the 20 amino acids in proteins (see Figure 14.6). Does this mean that the cell must produce 61 different tRNA species, each with a different anticodon? No. The cell gets by with about two-thirds of that number of tRNA species because the specificity for the base at the 3' end of the codon (and the 5' end of the anticodon) is not always strictly observed. This phenomenon, called *wobble*, allows the alanine codons GCA, GCC, and GCU, for example, all to be recognized by the same tRNA. Wobble is allowed in some matches but not in others; of most importance, it does not allow the ge-

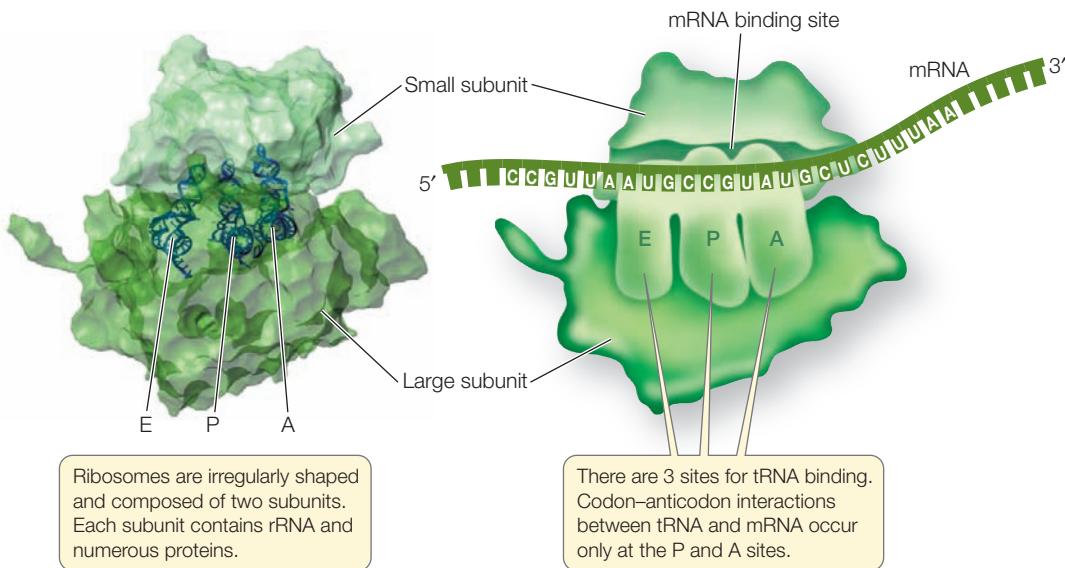
netic code to be ambiguous. That is, each mRNA codon binds to just one tRNA species, carrying a specific amino acid.

Activating enzymes link the right tRNAs and amino acids

The charging of each tRNA with its correct amino acid is achieved by a family of activating enzymes, known more formally as aminoacyl-tRNA synthases (Figure 14.13). Each activating enzyme is specific for one amino acid and for its corresponding tRNA. The enzyme has a three-part active site that recognizes three molecules: a specific amino acid, ATP, and a specific tRNA. Since tRNA has a complex three-dimensional structure, the activating enzyme recognizes a specific tRNA with a very low error rate. Remarkably, the error rate for amino acid recognition is also low, on the order of one in 1,000. Because the activating enzymes are so highly specific, the process of tRNA charging is sometimes called the *second genetic code*. Follow the events of activation in Figure 14.13.

A clever experiment by Seymour Benzer and his colleagues at Purdue University demonstrated the importance of specificity in the attachment of tRNA to its amino acid. In their laboratory, the amino acid cysteine, already properly attached to its tRNA, was chemically modified to become a different amino acid, alanine. Which component—the amino acid or the tRNA—would be recognized when this hybrid charged tRNA was put





14.14 Ribosome Structure Each ribosome consists of a large and a small subunit. The subunits remain separate when they are not in use for protein synthesis.

into a protein-synthesizing system? The answer was the tRNA. Everywhere in the synthesized protein where cysteine was supposed to be, alanine appeared instead. The cysteine-specific tRNA had delivered its cargo (alanine) to every mRNA codon for cysteine. This experiment showed that the protein synthesis machinery recognizes the anticodon of the charged tRNA, not the amino acid attached to it. If activating enzymes in nature did what Benzer did in the laboratory and charged tRNAs with the wrong amino acids, those amino acids would be inserted into proteins at inappropriate places, leading to alterations in protein shape and function and endangering cell life.

The ribosome is the workbench for translation

The **ribosome** is the molecular workbench where the task of translation is accomplished. Its structure enables it to hold mRNA and charged tRNAs in the right positions, thus allowing a polypeptide chain to be assembled efficiently. A given ribosome does not specifically produce just one kind of protein. A ribosome can use any mRNA and all species of charged tRNAs, and thus can be used to make many different polypeptide products. Ribosomes can be used over and over again, and there are thousands of them in a typical cell.

Although ribosomes are small in contrast to other cellular organelles, their mass of several million daltons makes them large in comparison with charged tRNAs. Each ribosome consists of two subunits, a large one and a small one (Figure 14.14). In eukaryotes, the large subunit consists of three different molecules of ribosomal RNA (rRNA) and 49 different protein molecules, arranged in a precise pattern. The small subunit consists of one rRNA molecule and 33 different protein molecules.

These two subunits and several dozen other molecules interact non-covalently, like a jigsaw puzzle. In fact, when hydrophobic interactions between the proteins and RNAs are disrupted, the ribosome falls apart. If the disrupting agent is removed, the complex structure self-assembles perfectly! When not active in the translation of mRNA, the ribosomes exist as two separate subunits.

The ribosomes of prokaryotes are somewhat smaller than those of eukaryotes, and their ribosomal proteins and RNAs are different. Mitochondria and chloroplasts also contain ribosomes, some of which are similar to those of prokaryotes (see Chapter 5).

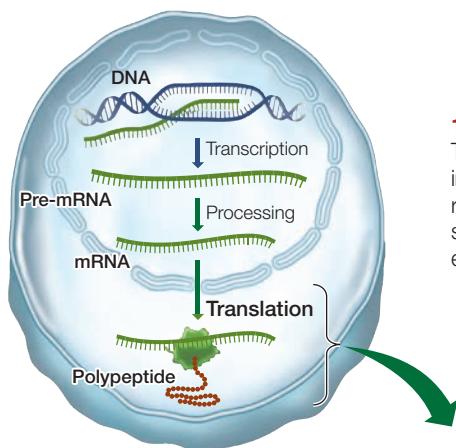
On the large subunit of the ribosome there are three sites to which a tRNA can bind, and these are designated A, P, and E (see Figure 14.14). The mRNA and ribosome move in relation to one another, and as they do so, a charged tRNA traverses these three sites in order:

- The *A (amino acid) site* is where the charged tRNA anticodon binds to the mRNA codon, thus lining up the correct amino acid to be added to the growing polypeptide chain.
- The *P (polypeptide) site* is where the tRNA adds its amino acid to the polypeptide chain.
- The *E (exit) site* is where the tRNA, having given up its amino acid, resides before being released from the ribosome and going back to the cytosol to pick up another amino acid and begin the process again.

The ribosome has a *accuracy function* that ensures that the mRNA-tRNA interactions are accurate; that is, that a charged tRNA with the correct anticodon (e.g., 3'-UAC-5') binds to the appropriate codon in mRNA (e.g., 5'-AUG-3'). When proper binding occurs, hydrogen bonds form between the base pairs. The rRNA of the small ribosomal subunit plays a role in validating the three-base-pair match. If hydrogen bonds have not formed between all three base pairs, the tRNA must be the wrong one for that mRNA codon, and that tRNA is ejected from the ribosome.

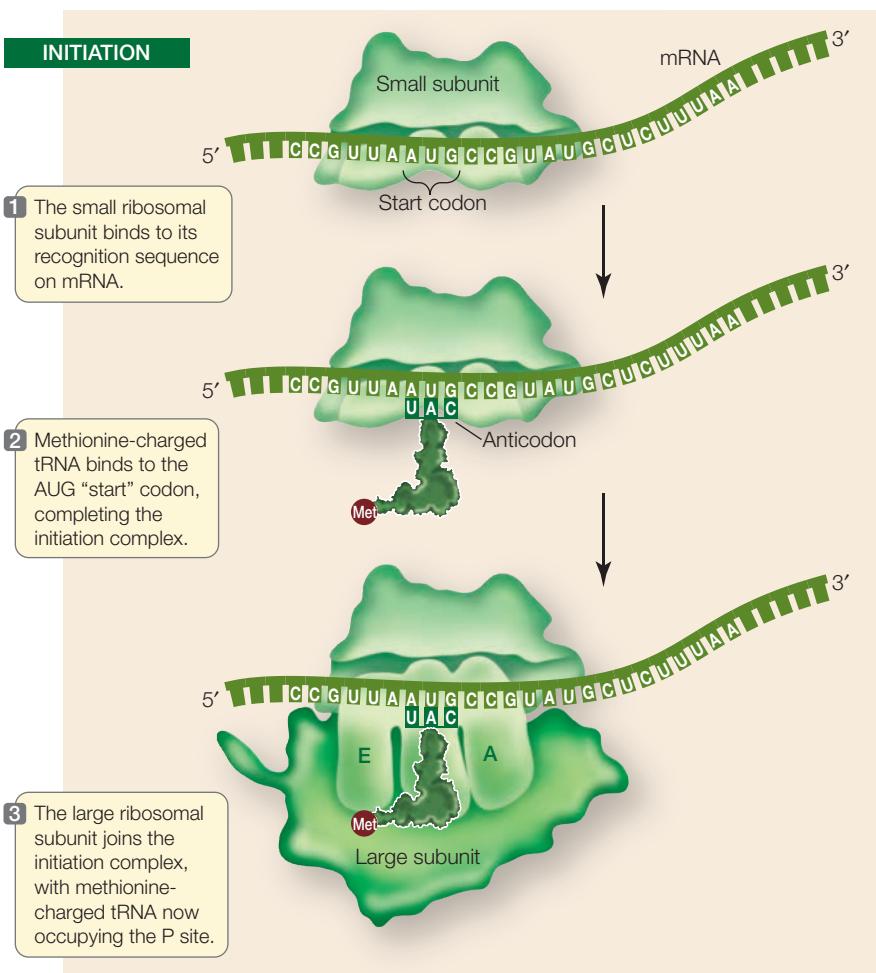
Translation takes place in three steps

Translation is the process by which the information in mRNA (derived from DNA) is used to specify and link a specific sequence of amino acids, producing a polypeptide. Like transcription, translation occurs in three steps: initiation, elongation, and termination.



14.15 The Initiation of Translation

Translation begins with the formation of an initiation complex. In prokaryotes, the small ribosomal subunit binds to the Shine-Dalgarno sequence to begin the process, while in eukaryotes, it binds to the 5' cap.



INITIATION The translation of mRNA begins with the formation of an **initiation complex**, which consists of a charged tRNA and a small ribosomal subunit, both bound to the mRNA (Figure 14.15).

In prokaryotes, the rRNA of the small ribosomal subunit first binds to a complementary ribosome binding site (known as the Shine–Dalgarno sequence) on the mRNA. This sequence is upstream of the actual start codon, but lines up the start codon so that it is will be adjacent to the P site of the large subunit.

Eukaryotes do this somewhat differently: the small ribosomal subunit binds to the 5' cap on the mRNA. After binding, the small subunit moves along the mRNA until it reaches the start codon.

Recall that the mRNA start codon in the genetic code is AUG (see Figure 14.6). The anticodon of a methionine-charged tRNA

binds to this start codon by complementary base pairing to complete the initiation complex. Thus the first amino acid in a polypeptide chain is always methionine. However, not all mature proteins have methionine as their N-terminal amino acid. In many cases, the initiator methionine is removed by an enzyme after translation.

After the methionine-charged tRNA has bound to the mRNA, the large subunit of the ribosome joins the complex. The methionine-charged tRNA now lies in the P site of the ribosome, and the A site is aligned with the second mRNA codon. These ingredients—mRNA, two ribosomal subunits, and methionine-charged tRNA—are put together properly by a group of proteins called *initiation factors*.

ELONGATION A charged tRNA whose anticodon is complementary to the second codon of the mRNA now enters the open A site of the large ribosomal subunit. The large subunit then catalyzes two reactions:

- It breaks the bond between the tRNA and its amino acid in the P site.
- It catalyzes the formation of a peptide bond between that amino acid and the one attached to the tRNA in the A site.

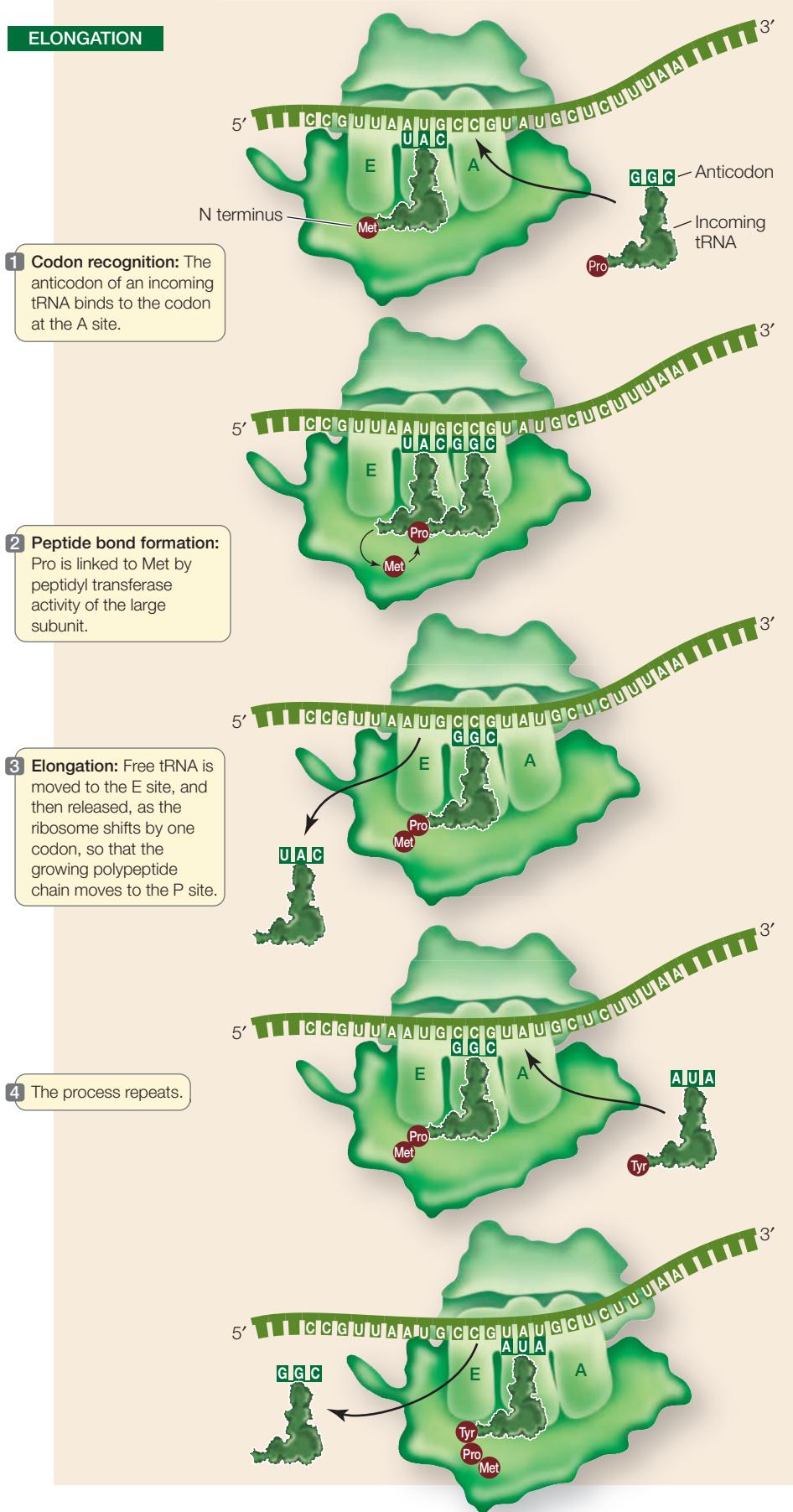
Because the large ribosomal subunit performs these two actions, it is said to have **peptidyl transferase** activity. In this way, methionine (the amino acid in the P site) becomes the N terminus of the new protein. The second amino acid is now bound to methionine, but remains attached to its tRNA at the A site.

How does the large ribosomal subunit catalyze this binding? Harry Noller and his colleagues at the University of California at Santa Cruz did a series of experiments and found that:

- If they removed almost all of the proteins from the large subunit, it still catalyzed peptide bond formation.
- If the rRNA was destroyed, so was peptidyl transferase activity.

Thus *rRNA is the catalyst*. The purification and crystallization of ribosomes has allowed scientists to examine their structure in detail, and the catalytic role of rRNA in peptidyl transferase activity has been confirmed. This supports the hypothesis that RNA, and catalytic RNA in particular, evolved before DNA (see Section 4.3).

After the first tRNA releases its methionine, it moves to the E site and is then dissociated from the ribosome, returning to the cytosol to become charged with another methionine. The second tRNA, now bearing a dipeptide (a two-amino-acid chain), is shifted to the P site as the ribosome moves one codon along the mRNA in the 5'-to-3' direction.

ELONGATION

14.16 The Elongation of Translation The polypeptide chain elongates as the mRNA is translated.

The elongation process continues, and the polypeptide chain grows, as these steps are repeated. Follow the process in **Figure 14.16**. All these steps are assisted by ribosomal proteins called *elongation factors*.

TERMINATION The elongation cycle ends, and translation is terminated, when a stop codon—UAA, UAG, or UGA—enters the A site (**Figure 14.17**). These codons do not correspond with any amino acids, nor do they bind any tRNAs. Rather, they bind a *protein release factor*, which allows hydrolysis of the bond between the polypeptide chain and the tRNA in the P site.

The newly completed polypeptide thereupon separates from the ribosome. Its C terminus is the last amino acid to join the chain. Its N terminus, at least initially, is methionine, as a consequence of the AUG start codon. In its amino acid sequence, it contains information specifying its conformation, as well as its ultimate cellular destination.

Table 14.2 summarizes the nucleic acid signals for initiation and termination of transcription and translation.

Polysome formation increases the rate of protein synthesis

Several ribosomes can work simultaneously at translating a single mRNA molecule, producing multiple polypeptides at the same time. As soon as the first ribosome has moved far enough from the site of translation initiation, a second initiation complex can form, then a third, and so on. An assemblage consisting of a strand of mRNA with its beadlike ribosomes and their growing polypeptide chains is called a **polyribosome**, or **polysome** (**Figure 14.18**). Cells that are actively synthesizing proteins contain large numbers of polysomes and few free ribosomes or ribosomal subunits.

A polysome is like a cafeteria line in which patrons follow one another, adding items to their trays. At any moment, the person at the start has a little food (a newly initiated protein); the person at the end has a complete meal (a completed protein). However, in the polysome cafeteria, everyone gets the same meal: many copies of the same protein are made from a single mRNA.

14.17 The Termination of Translation Translation terminates when the A site of the ribosome encounters a stop codon on the mRNA.

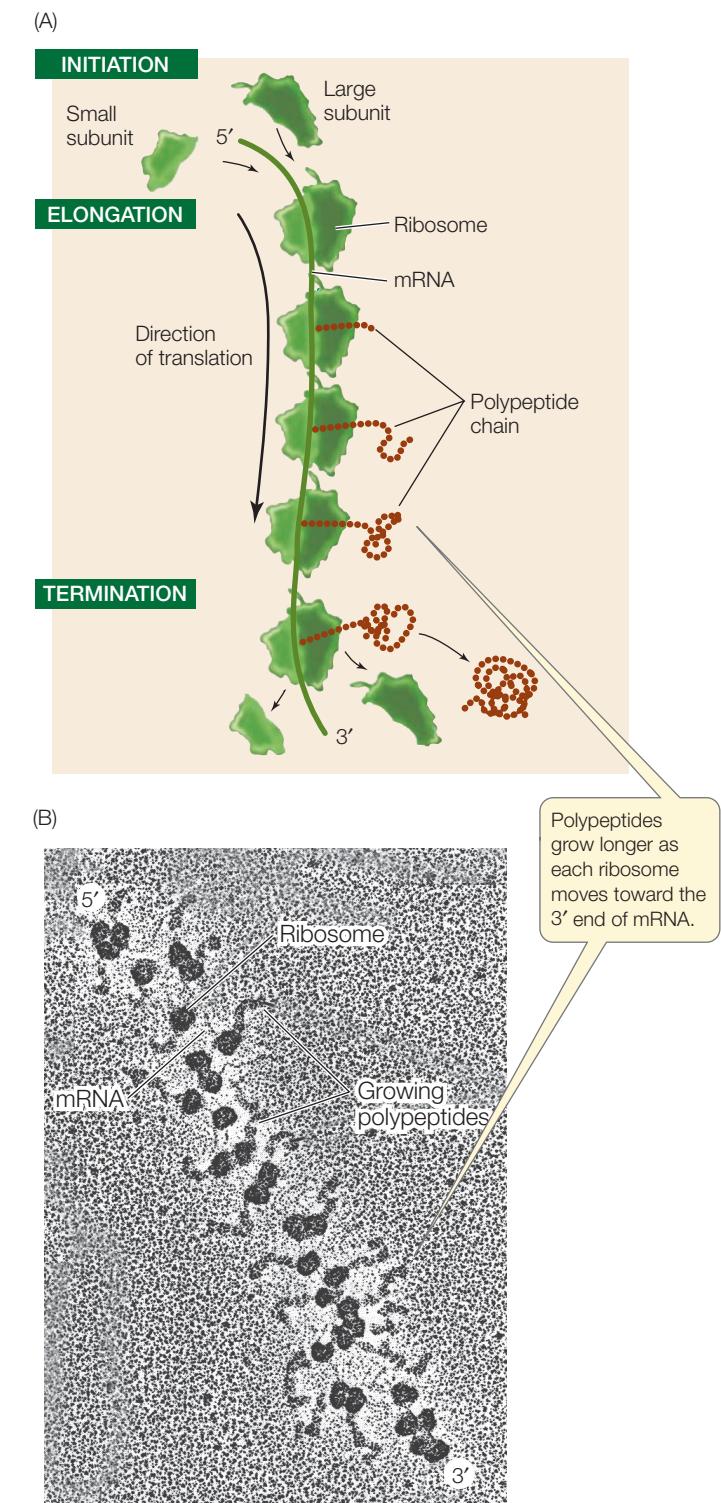
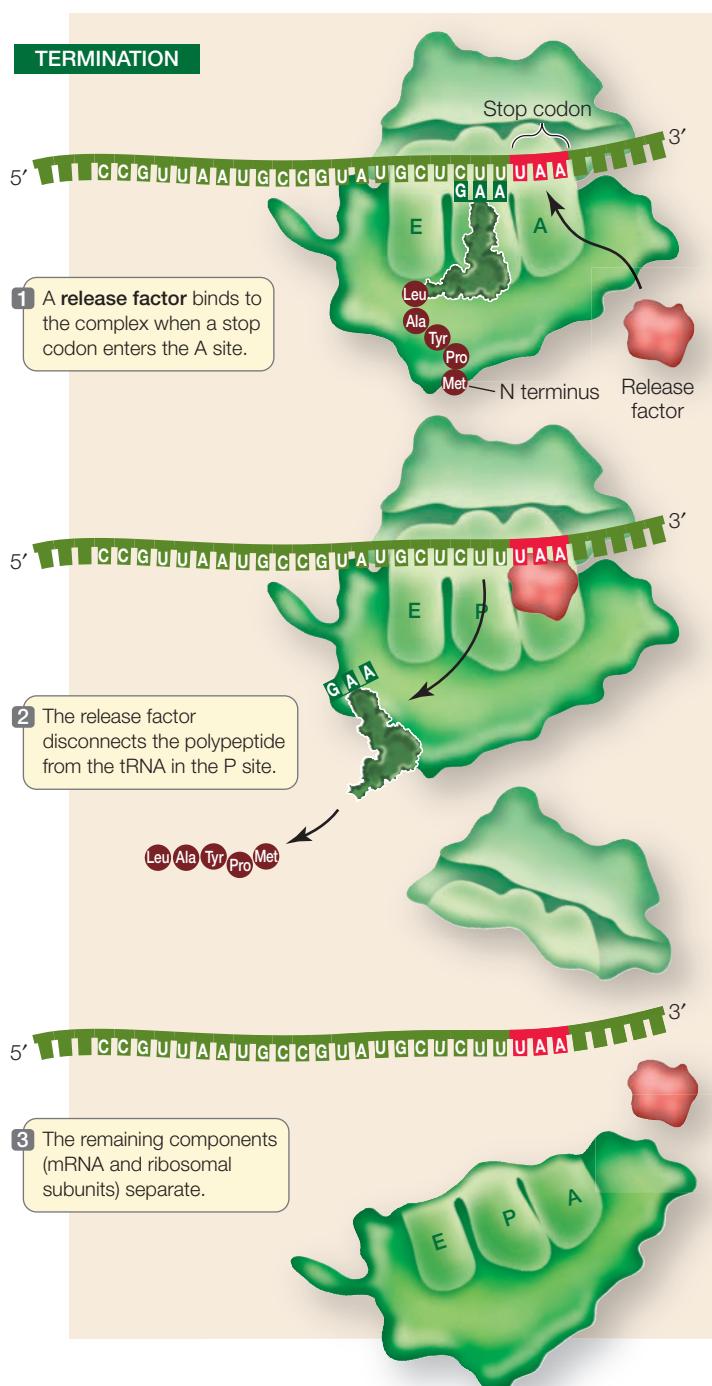


TABLE 14.2
Signals that Start and Stop Transcription and Translation

	TRANSCRIPTION	TRANSLATION
Initiation	Promoter DNA	AUG start codon in the mRNA
Termination	Terminator DNA	UAA, UAG, or UGA in the mRNA

14.18 A Polysome (A) A polysome consists of multiple ribosomes and their growing polypeptide chains moving along an mRNA molecule. (B) An electron micrograph of a polysome.

14.5 RECAP

A key step in protein synthesis is the attachment of an amino acid to its proper tRNA. This attachment is carried out by an activating enzyme. Translation of the genetic information from mRNA into protein occurs at the ribosome. Multiple ribosomes may act on a single mRNA to make multiple copies of the protein that it encodes.

- How is an amino acid attached to a specific tRNA, and why is the term “second genetic code” associated with this process? See pp. 304–305 and Figure 14.13
- Describe the events of initiation, elongation, and termination of translation. See pp. 306–308 and Figures 14.15–14.17

The polypeptide chain that is released from the ribosome is not necessarily a functional protein. Let’s look at some of the post-translational changes that can affect the fate and function of a polypeptide.

14.6 What Happens to Polypeptides after Translation?

The site of a polypeptide’s function may be far away from its point of synthesis in the cytoplasm. This is especially true for eukaryotes. The polypeptide may be moved into an organelle, or even out of the cell. In addition, polypeptides are often modified by the addition of new chemical groups that have func-

tional significance. In this section we examine these *posttranslational* aspects of protein synthesis.

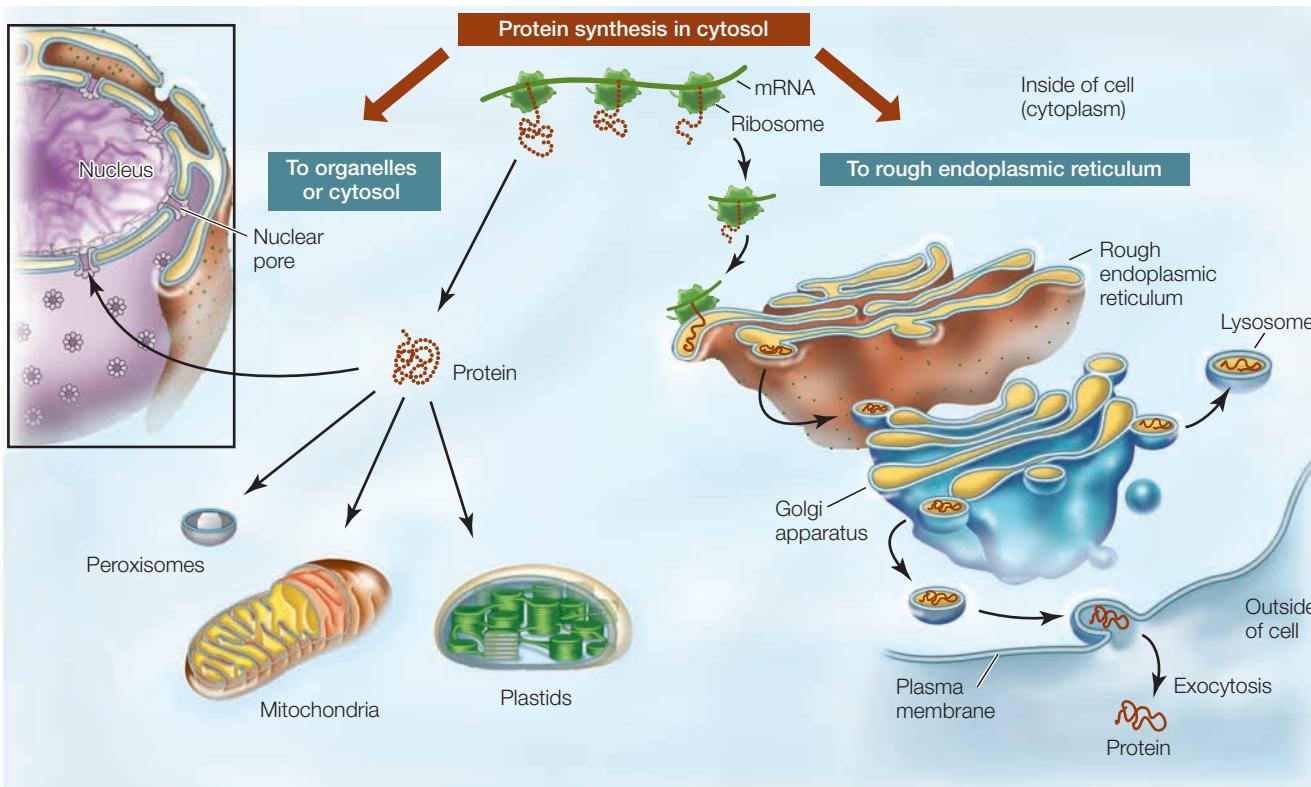
Signal sequences in proteins direct them to their cellular destinations

As a polypeptide chain emerges from the ribosome, it folds into its three-dimensional shape. As we described in Section 3.2, the polypeptide’s conformation is determined by the sequence of amino acids that make it up. Properties such as the polarity and charge of the R groups in the amino acids determine how they interact with each other in the folded molecule. Ultimately, a polypeptide’s conformation allows it to interact with other molecules in the cell, such as a substrate or another polypeptide. In addition to this structural information, the newly formed polypeptide can contain a **signal sequence**—an “address label” indicating where in the cell the polypeptide belongs.

Protein synthesis always begins on free ribosomes in the cytoplasm. But as a polypeptide chain is made, the information contained in its amino acid sequence gives it one of two sets of further instructions (Figure 14.19):

- “Complete translation and be released to an organelle, or remain in the cytosol.” Some proteins contain signal sequences that direct them to the nucleus, mitochondria, plastids, or per-

14.19 Destinations for Newly Translated Polypeptides in a Eukaryotic Cell Signal sequences on newly synthesized polypeptides bind to specific receptor proteins on the outer membranes of the organelles to which they are “addressed.” Once the protein has bound to it, the receptor forms a channel in the membrane, and the protein enters the organelle.



oxisomes. If they lack a signal sequence, they remain in the cytosol by default.

- “Stop translation, go to the endoplasmic reticulum, and finish synthesis there.” Other proteins contain a signal sequence that directs them to the endoplasmic reticulum (ER) before translation is complete. Such proteins may be retained in the lumen (the inside) of the ER, or be sent to the Golgi apparatus. From there, they may be sent to the lysosomes or the plasma membrane. Alternatively if they lack such specific instructions, they may be secreted from the cell via vesicles that fuse with the plasma membrane.

DESTINATION: NUCLEUS, MITOCHONDRIUM, OR CHLOROPLAST After translation, some folded polypeptides have a short exposed sequence of amino acids that acts like a postal “zip code,” directing them to an organelle. These signal (or localization) sequences are either at the N terminus or in the interior of the amino acid chain. For example, the following sequence directs a protein to the nucleus:

—Pro—Pro—Lys—Lys—Lys—Arg—Lys—Val—

A nuclear localization sequence would occur in histone proteins associated with nuclear DNA, but not in citric acid cycle enzymes, which are addressed to the mitochondria. Signal sequences for a particular organelle vary, so not all the polypeptides destined for the nucleus have the same signal sequence.

How do we know that the amino acid sequence shown above is the signal? To investigate this question, Stephen Dilworth and colleagues at the University of Cambridge injected cells with nuclear and cytoplasmic proteins (Figure 14.20). The experiments involved the nuclear protein nucleoplasmin, the cytoplasmic protein pyruvate kinase, the nuclear localization signal (see above) and a “mix-and-match” procedure. For example, the putative nuclear signal was removed from nucleoplasmin, which normally carries it, or attached to pyruvate kinase, which does not normally carry it. The result was that it did not matter where in the cell the protein normally resided. If it had the signal, it went to the nucleus and if it did not have the signal, it stayed in the cytoplasm.

A signal sequence binds to a specific receptor protein, appropriately called a **docking protein**, on the outer membrane of the appropriate organelle. Once the signal sequence has bound to it, the docking protein forms a channel in the membrane, allowing the signal-bearing protein to pass through the membrane and enter the organelle. In this process, the protein is usually unfolded by a chaperonin protein (see Figure 3.12) so that it can pass through the channel; then it refolds into its normal conformation.

DESTINATION: ENDOPLASMIC RETICULUM If a specific hydrophobic sequence of 15–30 amino acids occurs at the N terminus of an elongated polypeptide chain, the polypeptide is sent initially to the ER. Some proteins are retained in the ER, but most move on to the Golgi, where they can be modified for eventual transport to the lysosomes, the plasma membrane, or out of the cell. In the cytoplasm, before translation is finished and while the polypeptide is still attached to a ribosome, this signal sequence

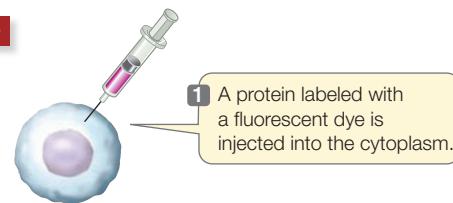
INVESTIGATING LIFE

14.20 Testing the Signal

A series of experiments were used to test whether the nuclear localization signal (NLS) sequence is all that is needed to direct a protein to the nucleus.

HYPOTHESIS A nuclear localization signal (NLS) is necessary for import of a protein into the nucleus.

METHOD



RESULTS

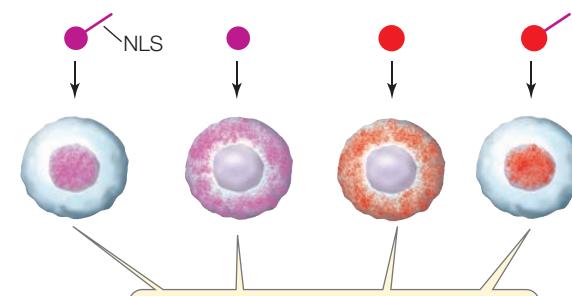
Injected protein:

Nucleoplasmin, a nuclear protein, with the NLS

Nucleoplasmin with the NLS removed

Pyruvate kinase, a cytoplasmic protein without the NLS

Pyruvate kinase, attached NLS



2 The distribution of the protein in the cell is observed with a fluorescence microscope.

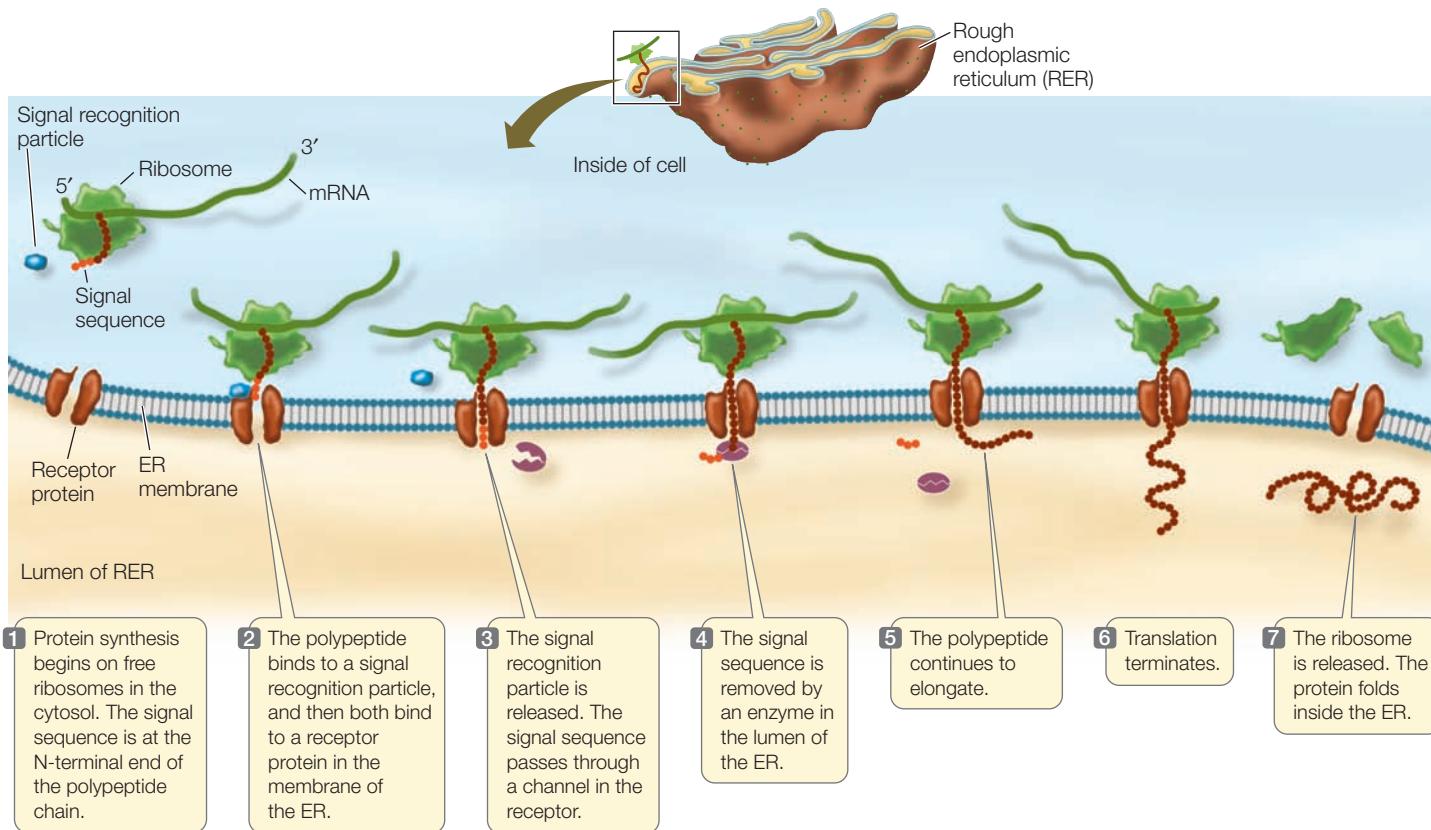
CONCLUSION

The NLS is essential for nuclear protein import and will direct a normally cytoplasmic protein to the nucleus.

FURTHER INVESTIGATION: How would you test for a chloroplast signal sequence?

Go to yourBioPortal.com for original citations, discussions, and relevant links for all INVESTIGATING LIFE figures.

binds to a **signal recognition particle** composed of protein and RNA (Figure 14.21). This binding blocks further protein synthesis until the ribosome becomes attached to a specific receptor protein in the membrane of the rough ER. Once again, the receptor protein is converted into a channel, through which the growing polypeptide passes. After the formation of the channel, protein synthesis resumes, and the chain grows longer until its sequence is completed. The elongating polypeptide may be retained in the ER membrane itself, or it may enter the



14.21 A Signal Sequence Moves a Polypeptide into the ER When a certain signal sequence of amino acids is present at the beginning of a polypeptide chain, the polypeptide will be taken into the endoplasmic reticulum (ER). The finished protein is thus segregated from the cytosol.

interior space—the lumen—of the ER. In either case, an enzyme in the lumen of the ER removes the signal sequence from the polypeptide chain.

If the finished protein enters the ER lumen, it can be transported to its appropriate location—to other cellular compartments or to the outside of the cell—via the ER and the Golgi apparatus, without mixing with other molecules in the cytoplasm.

After removal of the terminal signal sequence in the lumen of the ER, additional signals are needed to direct the protein to its destination. These signals are of two kinds:

- Some are sequences of amino acids that allow the protein's retention within the ER.
- Others are sugars, which are added in the Golgi apparatus. The resulting *glycoproteins* end up either at the plasma membrane or in a lysosome (or plant vacuole), or are secreted, depending on which sugars are added.

Proteins with no additional signals pass from the ER through the Golgi apparatus and are secreted from the cell.

The importance of signals is shown by Inclusion-cell (I-cell) disease, an inherited disease that causes death in early childhood. People with this disease have a mutation in the gene encoding a Golgi enzyme that adds targeting sugars to proteins destined for the lysosomes. As a result, enzymes that are essential for the hydrolysis of various macromolecules cannot reach

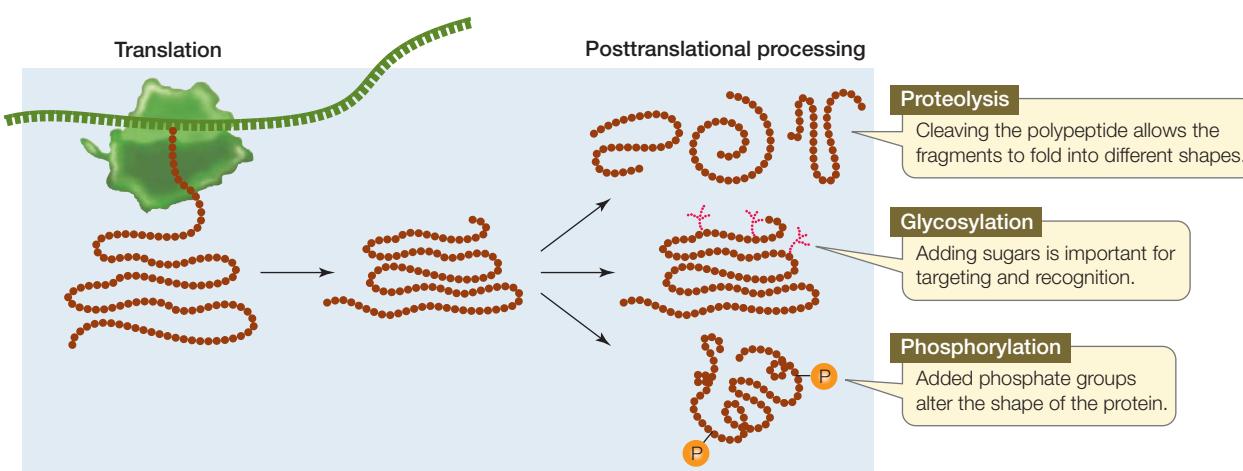
the lysosomes, where they are normally active. The macromolecules accumulate in the lysosomes, and this lack of cellular recycling has drastic effects, resulting in early death.

Many proteins are modified after translation

Most mature proteins are not identical to the polypeptide chains that are translated from mRNA on the ribosomes. Instead, most polypeptides are modified in any of a number of ways after translation (Figure 14.22). These modifications are essential to the final functioning of the protein.

- **Proteolysis** is the cutting of a polypeptide chain. Cleavage of the signal sequence from the growing polypeptide chain in the ER is an example of proteolysis; the protein might move back out of the ER through the membrane channel if the signal sequence were not cut off. Some proteins are actually made from *polyproteins* (long polypeptides) that are cut into final products by enzymes called *proteases*. These protease enzymes are essential to some viruses, including human immunodeficiency virus (HIV), because the large viral polyprotein cannot fold properly unless it is cut. Certain drugs used to treat acquired immune deficiency syndrome (AIDS) work by inhibiting the HIV protease, thereby preventing the formation of proteins needed for viral reproduction.

- **Glycosylation** is the addition of sugars to proteins to form glycoproteins. In both the ER and the Golgi apparatus, resident enzymes catalyze the addition of various sugars or short sugar chains to certain amino acid R groups on pro-



14.22 Posttranslational Modifications of Proteins Most polypeptides must be modified after translation in order to become functional proteins.

teins. One such type of “sugar coating” is essential for directing proteins to lysosomes, as mentioned above. Other types are important in the conformation of proteins and their recognition functions at the cell surface. Other attached sugars help to stabilize extracellular proteins, or proteins stored in vacuoles in plant seeds.

- **Phosphorylation** is the addition of phosphate groups to proteins, and is catalyzed by *protein kinases*. The charged phosphate groups change the conformation of a protein, often exposing the active site of an enzyme or the binding site for another protein. We have seen the role of phosphorylation in cell signaling (see Chapter 7).

14.6 RECAP

Signal sequences in polypeptides direct them to their appropriate destinations inside or outside the cell. Many polypeptides are modified after translation.

- How do signal sequences determine where a protein will go after it is made? See pp. 310–312 and Figure 14.21
- What are some ways in which posttranslational modifications alter protein structure and function? See pp. 312–313 and Figure 14.22

All of the processes we have just described result in a functional protein, but only if the amino acid sequence of that protein is correct. If the sequence is not correct, cellular dysfunction may result. Changes in the DNA—mutations—are a major source of errors in amino acid sequences. This is the subject of the next chapter.

CHAPTER SUMMARY

14.1 What Is the Evidence that Genes Code for Proteins?

- Beadle and Tatum’s experiments on metabolic enzymes in the bread mold *Neurospora* led to the **one-gene, one-enzyme hypothesis**. We now know that there is a **one-gene, one-polypeptide relationship**. Review Figure 14.1

14.2 How Does Information Flow from Genes to Proteins?

- The **central dogma** of molecular biology states that DNA encodes RNA, and RNA encodes proteins. Proteins do not encode proteins, RNA, or DNA.
- The process by which the information in DNA is copied to RNA is called **transcription**. The process by which a protein is built from the information in RNA is called **translation**. Review Figure 14.2, WEB ACTIVITY 14.1
- Certain RNA viruses are exceptions to the central dogma. These **retroviruses** synthesize DNA from RNA in a process called **reverse transcription**.
- The product of transcription is **messenger RNA (mRNA)**. **Transfer RNA (tRNA)** molecules are adapters that translate the

genetic information in the mRNA into a corresponding sequence of amino acids to produce a polypeptide.

14.3 How Is the Information Content in DNA Transcribed to Produce RNA?

- In a given gene, only one of the two strands of DNA (the **template strand**) acts as a template for transcription. **RNA polymerase** is the catalyst for transcription.
- RNA transcription from DNA proceeds in three steps: **initiation**, **elongation**, and **termination**. Review Figure 14.4, ANIMATED TUTORIAL 14.1
- Initiation requires a **promoter**, to which RNA polymerase binds. Part of each promoter is the **initiation site**, where transcription begins.
- Elongation of the RNA molecule proceeds from the 5' to 3' end.
- Particular base sequences specify termination, at which point transcription ends and the RNA transcript separates from the DNA template.

- The **genetic code** is a “language” of triplets of mRNA nucleotide bases (**codons**) corresponding to 20 specific amino acids; there are **start** and **stop codons** as well. The code is redundant (an amino acid may be represented by more than one codon), but not ambiguous (no single codon represents more than one amino acid). **Review Figures 14.5 and 14.6, ANIMATED TUTORIAL 14.2 AND WEB ACTIVITY 14.2**

14.4 How Is Eukaryotic DNA Transcribed and the RNA Processed?

- Unlike prokaryotes, where transcription and translation occur in the cytoplasm and are coupled, in eukaryotes transcription occurs in the nucleus and translation occurs later in the cytoplasm.
- The initial transcript of a eukaryotic protein-coding gene is modified with a **5' cap** and a **3' poly A sequence**. **Review Figure 14.10**
- Eukaryotic genes contain **introns**, which are noncoding sequences within the transcribed regions of genes.
- Pre-mRNA contains the introns. They are removed in the nucleus via **mRNA splicing** by the **small nuclear ribonucleoprotein particles**. Then the mRNA passes through the nuclear pore into the cytoplasm, where it is translated on the surfaces of **ribosomes**. **Review Figure 14.11, ANIMATED TUTORIAL 14.3**

14.5 How Is RNA Translated into Proteins?

SEE ANIMATED TUTORIAL 14.4

- During translation, amino acids are linked together in the order specified by the codons in the mRNA. This task is achieved by tRNAs, which bind to (are charged with) specific amino acids.
- Each tRNA species has an amino acid attachment site as well as an **anticodon** complementary to a specific mRNA codon. A specific activating enzyme charges each tRNA with its specific amino acid. **Review Figures 14.12 and 14.13**

- The **ribosome** is the molecular workbench where translation takes place. It has one large and one small subunit, both made of **ribosomal RNA** and proteins.
- Three sites on the large subunit of the ribosome interact with tRNA anticodons. The **A site** is where the charged tRNA anticodon binds to the mRNA codon; the **P site** is where the tRNA adds its amino acid to the growing polypeptide chain; and the **E site** is where the tRNA is released.
- Translation occurs in three steps: **initiation, elongation, and termination**.
- The **initiation complex** consists of tRNA bearing the first amino acid, the small ribosomal subunit, and mRNA. A specific complementary sequence on the small subunit rRNA binds to the transcription initiation site on the mRNA. **Review Figure 14.15**
- The growing polypeptide chain is elongated by the formation of peptide bonds between amino acids, catalyzed by the rRNA. **Review Figure 14.16**
- When a stop codon reaches the A site, it terminates translation by binding a release factor. **Review Figure 14.17**
- In a **polysome**, more than one ribosome moves along a strand of mRNA at one time. **Review Figure 14.18**

14.6 What Happens to Polypeptides after Translation?

- Signal sequences** of amino acids direct polypeptides to their cellular destinations. **Review Figure 14.19**
- Destinations in the cytoplasm include organelles, which proteins enter after being recognized and bound by surface receptors called **docking proteins**.
- Proteins “addressed” to the ER bind to a **signal recognition particle**. **Review Figure 14.21**
- Posttranslational modifications of polypeptides include **proteolysis**, in which a polypeptide is cut into smaller fragments; **glycosylation**, in which sugars are added; and **phosphorylation**, in which phosphate groups are added. **Review Figure 14.22**

SELF-QUIZ

- Which of the following is *not* a difference between RNA and DNA?
 - RNA has uracil; DNA has thymine.
 - RNA has ribose; DNA has deoxyribose.
 - RNA has five bases; DNA has four.
 - RNA is a single polynucleotide strand; DNA is a double strand.
 - RNA molecules are smaller than human chromosomal DNA molecules.
- Normally, *Neurospora* can synthesize all 20 amino acids. A certain strain of this mold cannot grow in minimal nutritional medium, but grows only when the amino acid leucine is added to the medium. This strain
 - is dependent on leucine for energy.
 - has a mutation affecting a biochemical pathway leading to the synthesis of carbohydrates.
 - has a mutation affecting the biochemical pathways leading to the synthesis of all 20 amino acids.
 - has a mutation affecting the biochemical pathway leading to the synthesis of leucine.
 - has a mutation affecting the biochemical pathways leading to the syntheses of 19 of the 20 amino acids.
- An mRNA has the sequence 5'-AUGAAUCCUAG-3'. What is the template DNA strand for this sequence?
 - 5'-TACTTTAGGATC-3'
 - 5'-ATGAAATCTAG-3'
 - 5'-GATCCTAAAGTA-3'
 - 5'-TACAAATCCTAG-3'
 - 5'-CTAGGATTCAT-3'
- The adapters that allow translation of the four-letter nucleic acid language into the 20-letter protein language are called
 - aminoacyl-tRNA synthetases.
 - transfer RNAs.
 - ribosomal RNAs.
 - messenger RNAs.
 - ribosomes.
- Which of the following does *not* occur after eukaryotic mRNA is transcribed?
 - Binding of RNA polymerase to the promoter
 - Capping of the 5' end
 - Addition of a poly A tail to the 3' end
 - Splicing out of the introns
 - Transport to the cytosol

6. Transcription
 - a. produces only mRNA.
 - b. requires ribosomes.
 - c. requires tRNAs.
 - d. produces RNA growing from the 5' end to the 3' end.
 - e. takes place only in eukaryotes.
 7. Which statement about translation is *not* true?
 - a. Translation is RNA-directed polypeptide synthesis.
 - b. An mRNA molecule can be translated by only one ribosome at a time.
 - c. The same genetic code operates in almost all organisms and organelles.
 - d. Any ribosome can be used in the translation of any mRNA.
 - e. There are both start and stop codons.
 8. Which statement about RNA is *not* true?
 - a. Transfer RNA functions in translation.
- b. Ribosomal RNA functions in translation.
 - c. RNAs are produced by transcription.
 - d. Messenger RNAs are produced on ribosomes.
 - e. DNA codes for mRNA, tRNA, and rRNA.
9. The genetic code
 - a. is different for prokaryotes and eukaryotes.
 - b. has changed during the course of recent evolution.
 - c. has 64 codons that code for amino acids.
 - d. has more than one codon for many amino acids
 - e. is ambiguous.
 10. Which statement about RNA splicing is *not* true?
 - a. It removes introns.
 - b. It is performed by small nuclear ribonucleoprotein particles (snRNPs).
 - c. It removes the introns at the ribosome.
 - d. It is usually directed by consensus sequences.
 - e. It shortens the RNA molecule.

FOR DISCUSSION

1. In rats, a gene 1,440 base pairs (bp) long codes for an enzyme made up of 192 amino acids. Discuss this apparent discrepancy. How long would the initial and final mRNA transcripts be?
2. Har Gobind Khorana at the University of Wisconsin synthesized artificial mRNAs such as poly CA (CACA ...) and poly CAA (CAACAAACAA ...). He found that poly CA codes for a polypeptide consisting of alternating threonine (Thr) and histidine (His) residues. There are two possible codons in poly CA, CAC and ACA. One of these must encode histidine and the other threonine—but which is which? The answer comes from results with poly CAA, which produces three different polypeptides: poly Thr, poly Gln (glutamine), and poly Asn (asparagine). (An artificial mRNA can be read, inefficiently, beginning at any point in the chain; there is no specific initiation signal. Thus poly CAA can be read as a polymer of CAA, of ACA, or of AAC.) Compare the results of the poly CA and poly CAA experiments, and determine which codon corresponds with threonine and which with histidine.
3. Look back at Question 2. Using the genetic code in Figure 14.6 as a guide, deduce what results Khorana would have obtained had he used poly UG and poly UGG as artificial mRNAs. In fact, very few such artificial mRNAs would have given useful results. For an example of what could happen, consider poly CG and poly CGG. If poly CG were the mRNA, a mixed polypeptide of arginine and alanine (Arg–Ala–Ala–Arg ...) would be obtained; poly CGG would give three polypeptides: poly Arg, poly Ala, and poly Gly (glycine). Can any codons be determined from only these data? Explain.
4. Errors in transcription occur about 100,000 times as often as errors in DNA replication. Why can this high rate be tolerated in RNA synthesis but not in DNA synthesis?

ADDITIONAL INVESTIGATION

Beadle and Tatum's experiments showed that a biochemical pathway could be deduced from mutant strains. In bacteria, the biosynthesis of the amino acid tryptophan (T) from the precursor chorismate (C) involves four intermediate chemical compounds, which we will call D, E, F, and G. Here are the phenotypes of various mutant strains. Each strain has a mutation in a gene for a different enzyme; + means growth with the indicated compound added to the medium, and 0 means no growth. Based on these data, order the compounds (C, D, E, F, G, and T) and enzymes (1, 2, 3, 4, and 5) in a biochemical pathway.

Mutant strain	Addition to medium					
	C	D	E	F	G	T
1	0	0	0	0	+	+
2	0	+	+	0	+	+
3	0	+	0	0	+	+
4	0	+	+	+	+	+
5	0	0	0	0	0	+

WORKING WITH DATA (GO TO yourBioPortal.com)

Deciphering the Genetic Code The identification of the first mRNA codons associated with specific amino acids was a landmark in molecular biology (Figure 14.5). In this hands-on

exercise, you will learn about the experimental protocol that Nirenberg and Matthei followed, using artificial mRNA, and analyze the results they obtained.

Baby 81

The tsunami of December 26, 2004, struck the coastal town of Kalmunai, Sri Lanka, with such force that 4-month-old Abilass Jeyarajah was torn from his mother's arms and swept away. Hours later, while his parents desperately searched the devastated town, their tiny son washed up on the beach a kilometer away, alive. A local schoolteacher found him and brought him to the hospital—the eighty-first patient admitted that day. The hospital was overwhelmed with 1,000 dead bodies, many of



them children. Since Abilass was alive and healthy, he was dubbed "Baby 81, the miracle baby" and became an instant celebrity among the staff as they went about their grim duties of caring for the injured and dying.

Meanwhile, the parents kept looking. Two days later, they met the schoolteacher, who told them about the baby he had found. Rushing to the hospital, the Jeyarajahs were elated to find their son, but were in for a rude shock. Eight other couples who had also lost infants were claiming Baby 81 as theirs. The baby remained in the hospital while the case went to court.

Judge M. P. Mohaideen faced a situation not unlike one faced by King Solomon 3,000 years ago, who was asked to decide which of two women was the mother of an infant. Solomon's method of determining parentage is told in a famous biblical passage—he ordered the baby cut in two, and the real mother indicated that she would rather give the baby away than have the baby killed. The Sri Lankan judge had a different method: he called in molecular biologists.

With 6 billion base pairs of DNA packaged in 46 chromosomes, each one of us is unique. Although our protein-coding sequences are similar (after all, our phenotypes are similar), only 1.5 percent of the DNA in the human genome actually codes for proteins. The eukaryotic genome contains many repeated sequences, and the repeat frequencies may differ between individuals, offering one way to differentiate one individual from another. A base pair at a particular site may also vary between individuals, due to DNA replication errors or random muta-

After the Tsunami In December of 2004, a tsunami originating in the Indian Ocean struck a broad region that encompassed many nations in Southeast Asia. The result was an unprecedented humanitarian disaster that left almost a quarter of a million people dead and many more homeless.



Baby 81 Abilass Jeyarajah survived the tsunami and was reunited with his parents by court order after DNA testing proved that he is indeed their son.

tions. Both of these types of differences are mutations, defined as inherited changes in DNA.

It is now possible to analyze these differences in DNA sequences (amplified by PCR) to identify people, in a process called DNA fingerprinting. The most common DNA fingerprinting technique used today involves the detection of variations in repeat sequences at different loci throughout the genome. When DNA samples from the nine sets of contesting parents were analyzed and compared with a sample from Baby 81, only one pair of parents carried sequences that were the same as those of the baby. On February 14, 2005, the judge ruled that the Jeyarajahs were the biological parents, and Baby 81 got his real name and parents back.

IN THIS CHAPTER we will discuss the nature and detection of mutations at the molecular and chromosomal levels. We will describe how abnormal proteins can cause human genetic diseases, and how these diseases and the alleles that produce them can be detected. Finally, we'll see how this knowledge of mutations has been applied in the development of new treatments.

CHAPTER OUTLINE

- 15.1 What Are Mutations?
- 15.2 How Are DNA Molecules and Mutations Analyzed?
- 15.3 How Do Defective Proteins Lead to Diseases?
- 15.4 What DNA Changes Lead to Genetic Diseases?
- 15.5 How Is Genetic Screening Used to Detect Diseases?
- 15.6 How Are Genetic Diseases Treated?

15.1 What Are Mutations?

In Chapter 12, we described mutations as inherited changes in genes, and we saw that different alleles may produce different phenotypes (short pea plants versus tall, for example). Now that we understand the chemical nature of genes and how they are expressed as phenotypes (in particular, proteins) we can return to the concept of mutations for a more specific definition. We can now state that mutations are changes in the nucleotide sequence of DNA that are passed on from one cell, or organism, to another.

As an example of just one cause of mutations, recall from Chapter 13 that DNA polymerases make errors. Repair systems such as proofreading are in place to correct them. But some errors escape being corrected and are passed on to the daughter cells.

Mutations in multicellular organisms can be divided into two types:

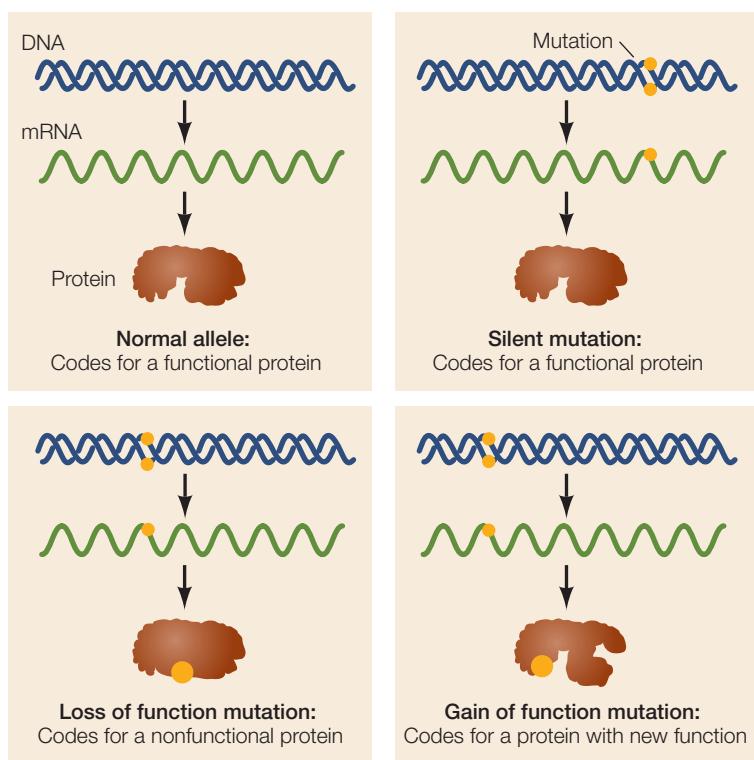
- **Somatic mutations** are those that occur in somatic (body) cells. These mutations are passed on to the daughter cells during mitosis, and to the offspring of those cells in turn, but are not passed on to sexually produced offspring. For example, a mutation in a single human skin cell could result in a patch of skin cells that all have the same mutation, but it would not be passed on to the person's children.
- **Germ line mutations** are those that occur in the cells of the germ line—the specialized cells that give rise to gametes. A gamete with the mutation passes it on to a new organism at fertilization.

In either case, the mutations may or may not have phenotypic effects.

Mutations have different phenotypic effects

Phenotypically, we can understand mutations in terms of their effects on proteins and their function (**Figure 15.1**).

- **Silent mutations** do not affect protein function. They can be mutations in noncoding DNA, such as the repeat sequences that were used to identify Baby 81 in the opening story of this chapter. Or they can be in the coding portion of DNA but not have any effect on the protein.
- **Loss of function mutations** affect protein function. These mutations may lead to nonfunctional proteins that no longer work as structural proteins or enzymes. They almost



15.1 Mutation and Phenotype Mutations may or may not affect the protein phenotype.

always show recessive inheritance in a diploid organism, because the presence of one wild-type allele will usually result in sufficient functional protein for the cell. For example, the familiar wrinkled-seed allele in pea plants, originally studied by Mendel (see Figure 12.3), is due to a mutation in the gene *SBE1* (starch branching enzyme). Normally the protein made by this gene catalyzes the branching of starch as seeds develop. In the mutant, the *SBE1* protein is not functional and that leads to osmotic changes, causing the wrinkled appearance.

- A **gain of function mutation** leads to a protein with an altered function. This kind of mutation usually shows dominant inheritance, because the presence of the wild-type allele does not prevent the mutant allele from functioning. This is common in cancer. For example, a receptor for a growth factor normally requires binding of the growth factor (the ligand) to activate the cell division cycle. Some cancers are caused by mutations in genes coding for these receptors such that they no longer require stimulation by their particular ligands. The mutant receptors are “always on,” leading to the unrestrained cell proliferation that is characteristic of cancer cells.
- **Conditional mutations** cause their phenotypes only under certain *restrictive* conditions. They are not detectable under other, *permissive* conditions. Many conditional mutants are temperature-sensitive; that is, they show the altered phenotype only at a certain temperature (recall the rabbit in Figure 12.11). The mutant allele in such an organism may code

for an enzyme with an unstable tertiary structure that is altered at the restrictive temperature.

All mutations are alterations in the nucleotide sequence of DNA. At the molecular level, we can divide mutations into two categories:

- A **point mutation** results from the gain, loss, or substitution of a single nucleotide. After DNA replication, the altered nucleotide becomes a mutant base pair. If a point mutation occurs within a gene (rather than in a noncoding DNA sequence), then one allele of that gene (usually dominant) becomes another allele (usually recessive).
- **Chromosomal mutations** are more extensive than point mutations. They may change the position or orientation of a DNA segment without actually removing any genetic information, or they may cause a segment of DNA to be duplicated or irretrievably lost.

Point mutations change single nucleotides

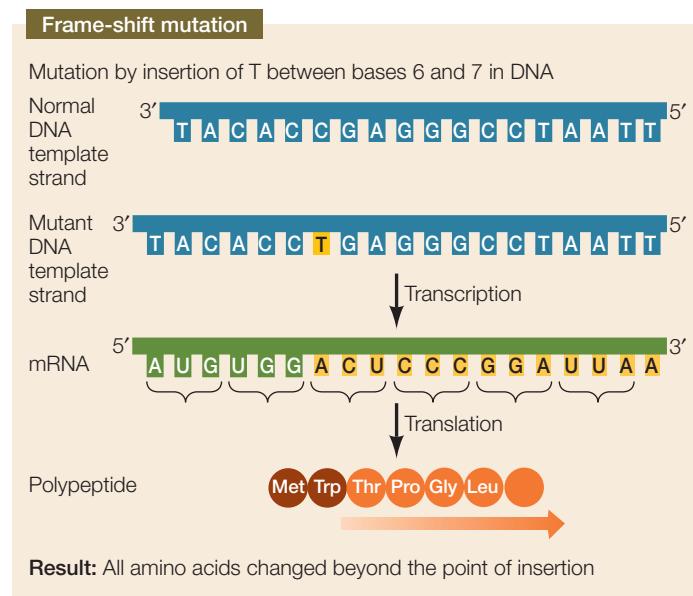
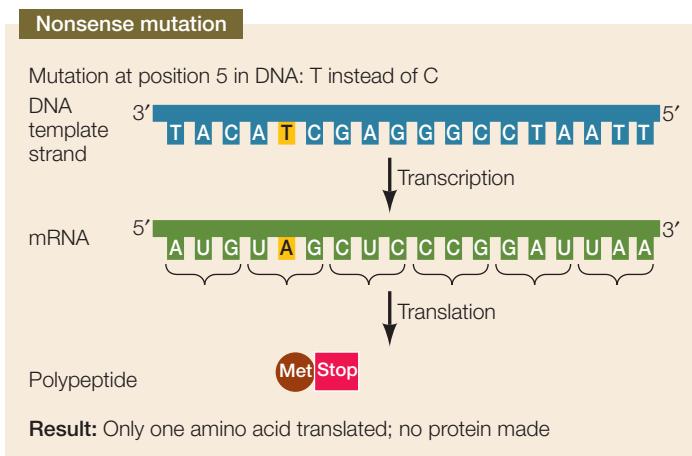
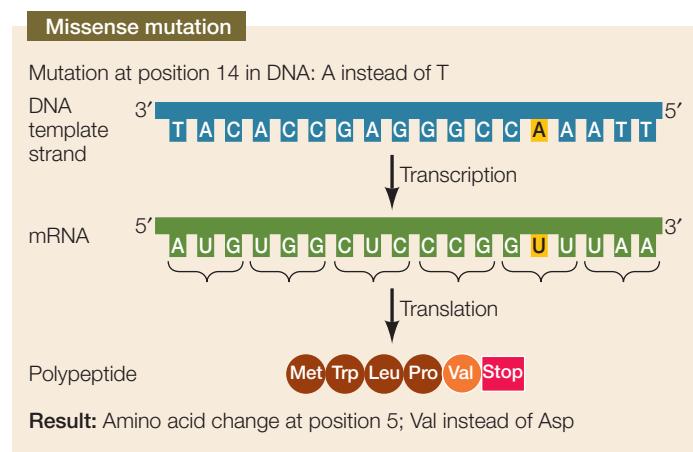
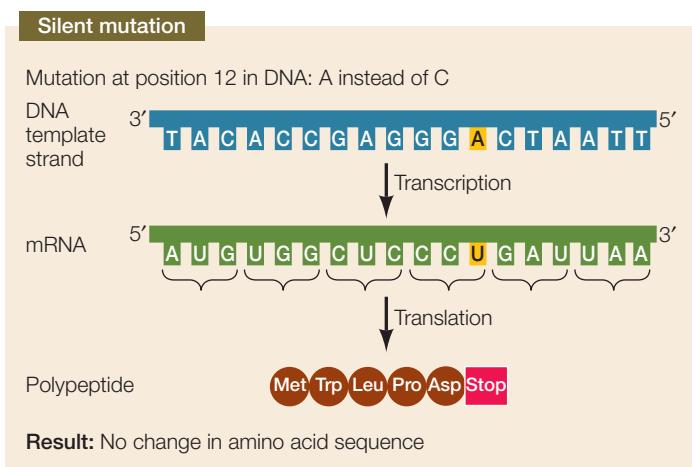
Point mutations result from the addition or subtraction of a nucleotide base, or the substitution of one base for another. Point mutations can arise due to errors in DNA replication that are not corrected during proofreading, or they may be caused by environmental **mutagens** (substances that cause mutations, such as radiation or certain chemicals).

Point mutations in the coding regions of DNA usually result in changes in the mRNA, but changes in the mRNA may or may not result in changes in the protein. Silent mutations by definition have no effect on the protein. Missense and nonsense mutations result in changes in the protein, some of them drastic (Figure 15.2).

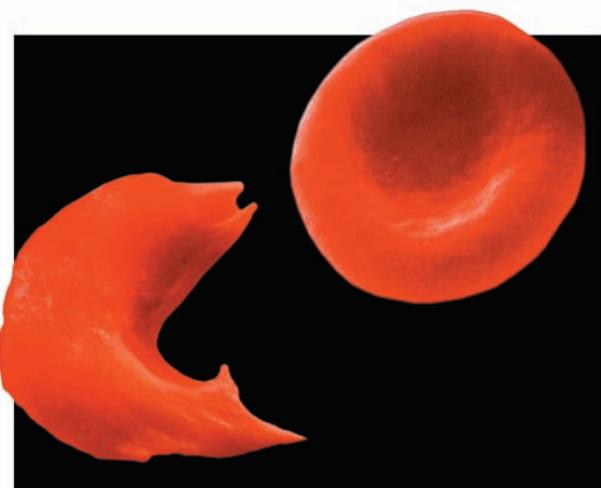
SILENT MUTATIONS Silent mutations have no effect on amino acid sequences. This is because they are often found in noncoding DNA. Also, because of the redundancy of the genetic code, a base substitution in a coding region will not always cause a change in the amino acid sequence when the altered mRNA is translated. Silent mutations are quite common, and they result in genetic diversity that is not expressed as phenotypic differences.

MISSENSE MUTATIONS Some base substitutions change the genetic code such that one amino acid substitutes for another in a protein. These changes are called **missense mutations**. A specific example of a missense mutation is the one that causes sickle-cell disease, a serious heritable blood disorder. The disease occurs in people who carry two copies of the sickle allele of the gene for human β -globin (a subunit of hemoglobin, the protein in human blood that carries oxygen). The sickle allele differs from the normal allele by one base pair, resulting in a polypeptide that differs by one amino acid from the normal protein. Individuals who are homozygous for this recessive allele have defective, sickle-shaped red blood cells (Figure 15.3).

A missense mutation may result in a defective protein, but often it has no effect on the protein’s function. For example, a



15.2 Point Mutations When they occur in the coding regions of proteins, single-base pair changes can cause missense, nonsense, or frame-shift mutations. Some of these mutations are silent, while others affect the protein's amino acid sequence.



15.3 Sickled and Normal Red Blood Cells The misshapen red blood cell on the left is caused by a missense mutation and an incorrect amino acid in one of the two polypeptides of hemoglobin.

hydrophilic amino acid may be substituted for another hydrophilic amino acid, so that the shape of the protein is unchanged. Or a missense mutation might reduce the functional efficiency of a protein rather than completely inactivating it. Therefore, individuals homozygous for a missense mutation in a protein essential for life may survive if enough of the protein's function is retained.

In some cases, a gain of function missense mutation occurs. An example is a mutation in the human *TP53* gene, which codes for a tumor suppressor; that is, the TP53 protein normally functions to inhibit the cell cycle. Certain mutations of the *TP53* gene cause this protein to no longer inhibit cell division, but to promote it and prevent programmed cell death. So a TP53 protein mutated in this way has a gain of oncogenic (cancer-causing) function.

NONSENSE MUTATIONS A **nonsense mutation** involves a base substitution that causes a stop codon (for translation) to form somewhere in the mRNA. A nonsense mutation results in a shortened protein, since translation does not proceed beyond

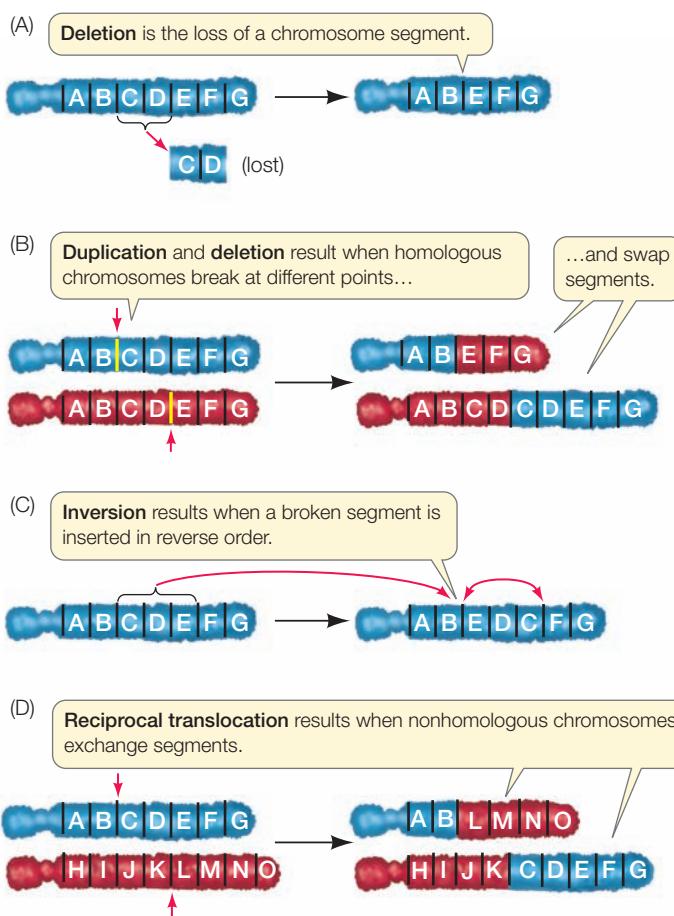
the point where the mutation occurred. For example, a common mutation causing thalassemia (another blood disorder affecting hemoglobin) in Mediterranean populations is a nonsense mutation that drastically shortens the α -globin subunit. Shortened proteins are usually not functional; however, if the nonsense mutation occurs near the 3' end of the gene, it may have no effect on function.

FRAME-SHIFT MUTATIONS Not all point mutations are base substitutions. Single or double bases may be inserted into or deleted from DNA. Such mutations in coding sequences are known as **frame-shift mutations** because they interfere with the translation of the genetic message by throwing it out of register. Think again of codons as three-letter words, each corresponding to a particular amino acid. Translation proceeds codon by codon; if a base is added to the mRNA or subtracted from it, translation proceeds perfectly until it comes to the one-base insertion or deletion. From that point on, the three-letter words in the genetic message are one letter out of register. In other words, such mutations shift the “reading frame” of the message. Frame-shift mutations almost always lead to the production of nonfunctional proteins.

Chromosomal mutations are extensive changes in the genetic material

Changes in single nucleotides are not the most dramatic changes that can occur in the genetic material. Whole DNA molecules can break and rejoin, grossly disrupting the sequence of genetic information. There are four types of such chromosomal mutations: *deletions*, *duplications*, *inversions*, and *translocations*. These mutations can be caused by severe damage to chromosomes resulting from mutagens or by drastic errors in chromosome replication.

- **Deletions** result from the removal of part of the genetic material (**Figure 15.4A**). Like frame-shift point mutations, their consequences can be severe unless they affect noncoding DNA or unnecessary genes, or are masked by the presence of normal alleles of the deleted genes in the same cell. It is easy to imagine one mechanism that could produce deletions: a DNA molecule might break at two points and the two end pieces might rejoin, leaving out the DNA between the breaks.
- **Duplications** can be produced at the same time as deletions (**Figure 15.4B**). A duplication would arise if homologous chromosomes broke at different positions and then reconnected to the wrong partners. One of the two chromosomes produced by this mechanism would lack a segment of DNA (it would have a deletion), and the other would have two copies (a duplication) of the segment that was deleted from the first chromosome.
- **Inversions** can also result from breaking and rejoining of chromosomes. A segment of DNA may be removed and reinserted into the same location in the chromosome, but “flipped” end over end so that it runs in the opposite direction (**Figure 15.4C**). If the break site includes part of a DNA



15.4 Chromosomal Mutations Chromosomes may break during replication, and parts of chromosomes may then rejoin incorrectly. The letters on these chromosome illustrations represent large segments of the chromosomes. Each segment may include anywhere from zero to hundreds or thousands of genes.

segment that codes for a protein, the resulting protein will be drastically altered and almost certainly nonfunctional.

- **Translocations** result when a segment of a chromosome breaks off and is inserted into a different chromosome. Translocations may involve reciprocal exchanges of chromosome segments, as in **Figure 15.4D**. Translocations often lead to duplications and deletions and may result in sterility if normal chromosome pairing in meiosis cannot occur.

Mutations can be spontaneous or induced

It is useful to distinguish two types of mutations in terms of their causes:

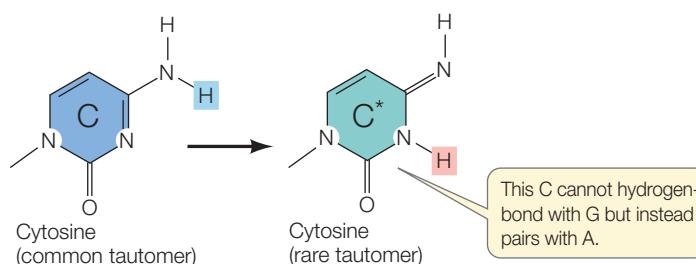
- **Spontaneous mutations** are permanent changes in the genetic material that occur without any outside influence. In other words, they occur simply because cellular processes are imperfect.
- **Induced mutations** occur when some agent from outside the cell—a mutagen—causes a permanent change in DNA.

Spontaneous mutations may occur by several mechanisms:

- The four nucleotide bases of DNA can have different structures. Each can exist in two different forms (called *tautomers*), one of which is common and one rare. When a base temporarily forms its rare tautomer, it can pair with the wrong base. For example, C normally pairs with G, but if C is in its rare tautomer at the time of DNA replication, it pairs with (and DNA polymerase will insert) an A. The result is a point mutation: G → A (**Figure 15.5A and C**).
- Bases in DNA may change because of a chemical reaction—for example, loss of an amino group in cytosine (a reaction called *deamination*). If this occurs in a DNA molecule, the error will usually be repaired. However, since the repair mechanism is not perfect, the altered nucleotide will sometimes remain during replication. Then, DNA polymerase will add an A (which base-pairs with U) instead of G (which normally pairs with C).
- DNA polymerase can make errors in replication (see Section 13.4)—for example, inserting a T opposite a G. Most of these errors are repaired by the proofreading function of the replication complex, but some errors escape detection and become permanent.

15.5 Spontaneous and Induced Mutations (A) All four nitrogenous bases in DNA exist in both a prevalent (common) form and a rare form. When a base spontaneously forms its rare tautomer, it can pair with a different base. (B) Mutagens such as nitrous acid can induce changes in the bases. (C) The results of both spontaneous and induced mutations are permanent changes in the DNA sequence following replication.

(A) A spontaneous mutation

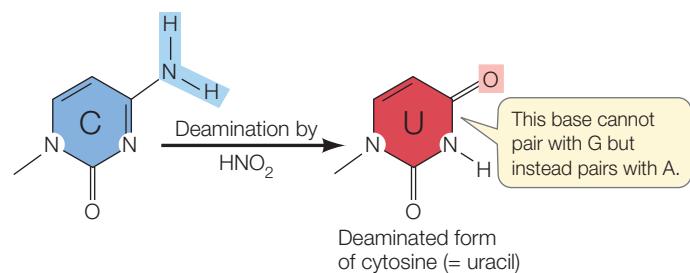


- Meiosis is not perfect. Nondisjunction—failure of homologous chromosomes to separate during meiosis—can occur, leading to one too many or one too few chromosomes (aneuploidy; see Figure 11.21). Random chromosome breakage and rejoining can produce deletions, duplications, inversions, or translocations.

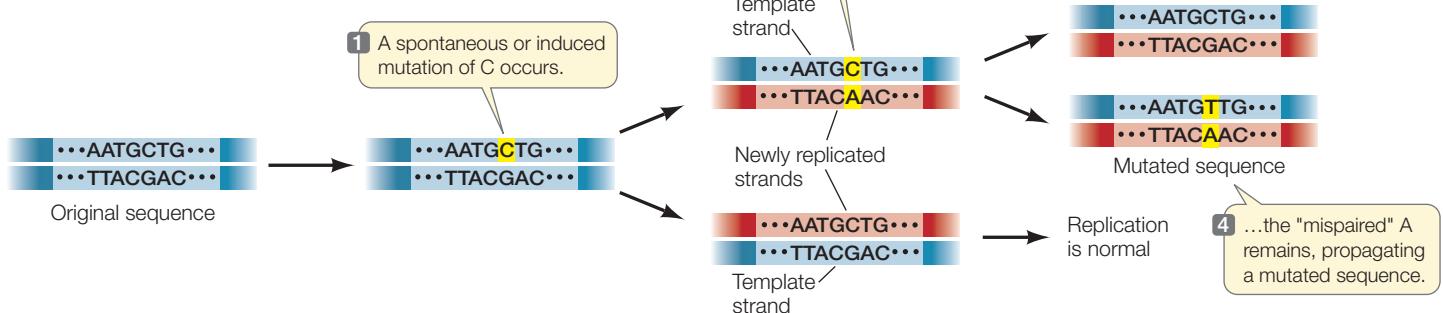
Induced mutations result from alterations of DNA by mutagens:

- Some chemicals can alter the nucleotide bases. For example, nitrous acid (HNO_2) and similar molecules can react with cytosine and convert it to uracil by deamination. More specifically, they convert an amino group on the cytosine ($-\text{NH}_2$) into a keto group ($-\text{C}=\text{O}$). This alteration has the same result as spontaneous deamination: instead of a G, DNA polymerase inserts an A (**Figure 15.5B and C**).
- Some chemicals add groups to the bases. For instance, benzoquinone, a component of cigarette smoke, adds a large chemical group to guanine, making it unavailable for base pairing. When DNA polymerase reaches such a modified guanine, it inserts any one of the four bases; of course, three-fourths of the time the inserted base will not be cytosine, and a mutation results.
- Radiation damages the genetic material. Radiation can damage DNA in two ways. First, ionizing radiation (including X rays, gamma rays, and radiation from unstable isotopes) produces highly reactive chemicals called *free radicals*. Free radicals can change bases in DNA to forms that are not recognized by DNA polymerase. Ionizing radiation can also

(B) An induced mutation



(C) The consequences of either mutation



break the sugar–phosphate backbone of DNA, causing chromosomal abnormalities. Second, ultraviolet radiation (from the sun or a tanning lamp) can damage DNA in another way. It is absorbed by thymine, causing it to form covalent bonds with adjacent bases. This, too, plays havoc with DNA replication by distorting the double helix.

Some base pairs are more vulnerable than others to mutation

DNA sequencing has revealed that mutations occur most often at certain base pairs. These “hot spots” are often located where cytosine has been methylated to 5-methylcytosine.

As we discussed above, unmethylated cytosine can lose its amino group, either spontaneously or because of a chemical mutagen, to form uracil (see Figure 15.5B). This type of error is usually detected by the cell and repaired. The DNA repair mechanism recognizes uracil as inappropriate for DNA (since uracil occurs only in RNA) and replaces it with cytosine.

However, when 5-methylcytosine loses its amino group, the product is thymine, a natural base for DNA. The DNA repair mechanism ignores this thymine (Figure 15.6). During replication, however, the mismatch repair mechanism recognizes that G-T is a mismatched pair, although it cannot tell which base was incorrectly inserted into the sequence. So half of the time it matches a new C to the G, but the other half of the time it matches a new A to the T, resulting in a mutation. It is not surprising that 5-methylcytosine residues are hot spots for mutation.

Mutagens can be natural or artificial

Many people associate mutagens with materials made by humans, but just as there are many human-made chemicals that cause mutations, there are also many mutagenic substances that occur naturally. Plants (and to a lesser extent animals) make thousands of small molecules that they use for their own purposes, such as defense against pathogens (see Chapter 39). Some of these are mutagenic and potentially carcinogenic. Examples of human-made mutagens are nitrites, which are used to preserve meats. Once in mammals, nitrites get converted by the smooth endoplasmic reticulum (ER) to nitrosamines, which are strongly mutagenic because they cause deamination of cytosine (see above). An example of a naturally occurring mutagen is aflatoxin, which is made by the mold *Aspergillus*. When mammals ingest the mold, the aflatoxin is converted by the ER into a product that, like benzopyrene from cigarette smoke, binds to guanine; this also causes mutations.

15.6 5-Methylcytosine in DNA Is a “Hot Spot” for Mutations

If cytosine has been methylated to 5-methylcytosine, the mutation is unlikely to be repaired and a C-G base pair is replaced with a T-A pair.

Radiation can also be human-made or natural. Some of the isotopes made in nuclear reactors and nuclear bomb explosions are certainly harmful. For example, extensive studies have shown increased mutations in the survivors of the atom bombs dropped on Japan in 1945. You probably know that natural ultraviolet radiation in sunlight also causes mutations, in this case by affecting thymine and, to a lesser extent, other bases in DNA.

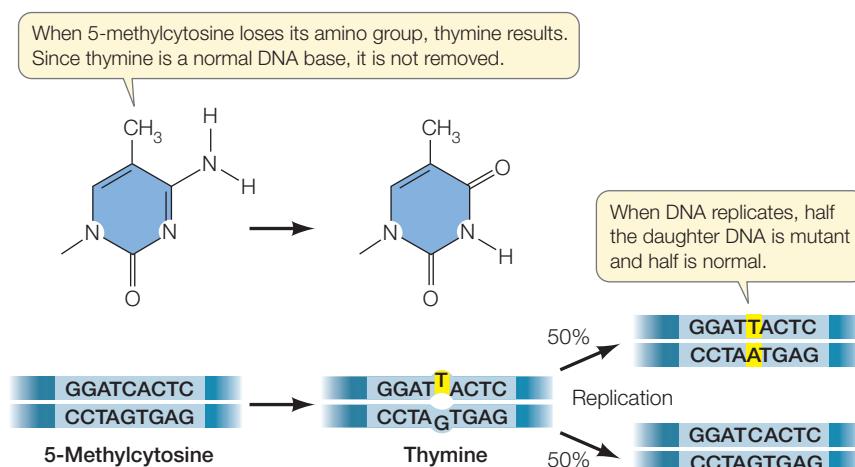
Biochemists have estimated how much DNA damage occurs in the human genome under normal circumstances: among the genome’s 2.3 billion base pairs there are about 16,000 DNA-damaging events per cell per day, of which 80 percent are repaired.

Mutations have both benefits and costs

What is the overall effect of mutation? For an organism, there are benefits and costs.

- *Mutations are the raw material of evolution.* Without mutation, there would be no evolution. As we will see in Part Seven of this book, mutation alone does not drive evolution, but it provides the genetic diversity that makes natural selection possible. This diversity can be beneficial in two ways. First, a mutation in somatic cells may benefit the organism immediately. Second, a mutation in germ line cells may have no immediate selective advantage to the organism but may cause a phenotypic change in offspring. If the environment changes in a later generation, that mutation may be advantageous and thus selected for under these conditions.
- *Germ line and somatic mutations can be harmful.* Mutations in germ line cells that get carried to the next generation are often deleterious, especially if the offspring are homozygous for a harmful recessive allele. In their extreme form, such mutations produce phenotypes that are lethal. Lethal mutations can kill an organism during early development, or the organism may die before maturity and reproduction.

In Chapter 11 we described how genetic changes in somatic cells can lead to cancer. Typically these are mutations in oncogenes (the “gas pedal”) that result in the stimulation of cell division, or mutations in tumor suppressor genes (the “brakes”) that result in a lack of inhibition of cell division. These muta-



tions can occur by either spontaneous or induced mutagenesis. While spontaneous mutagenesis is not in our control, we can certainly try to avoid mutagenic substances and radiation. Not surprisingly, many things that cause cancer (carcinogens) are also mutagens. A good example is benzopyrene (discussed above), which is found in coal tar, car exhaust fumes, and charbroiled foods, as well as in cigarette smoke.

A major environmental issue is the effect of both human-made and natural mutagens on public health. Identifying mutagens to which people are exposed, and estimating their risk for both mutagenesis and carcinogenesis, is a major public policy goal. Here are two recent examples:

- The Montreal Protocol is the only international environmental agreement signed and adhered to by all nations. It bans chlorofluorocarbons and other substances that cause depletion of the ozone layer in the upper atmosphere of Earth. Such depletion can result in increased ultraviolet radiation reaching Earth's surface. This would cause more somatic mutations which lead to skin cancer.
- Bans on cigarette smoking have rapidly spread throughout the world. Cigarette smoking causes cancer due to increased exposure of somatic cells in the lungs and throat to benzopyrene and other carcinogens.

15.1 RECAP

Mutations are alterations in the nucleotide sequence of DNA. They may be changes in single nucleotides or extensive rearrangements of chromosomes. If they occur in somatic cells, they will be passed on to daughter cells; if they occur in germ line cells, they will be passed on to offspring.

- What are the various kinds of point mutations? See pp. 318–320 and Figure 15.2
- What distinguishes the various kinds of chromosomal mutations: deletions, duplications, inversions, and translocations? See p. 320 and Figure 15.4
- Explain the difference between spontaneous and induced mutagenesis. Give an example of each. See pp. 320–322 and Figure 15.5
- Why do many mutations involve G-C base pairs? See p. 322 and Figure 15.6

We have seen that there are many different ways in which DNA can be altered, in terms of both the types of changes and the mechanisms by which they occur. We turn now to the ways that biologists detect mutations in DNA.

15.2 How Are DNA Molecules and Mutations Analyzed?

Once biologists understood the connections between phenotype and proteins, and between genes and DNA, they were faced with the important challenge of precisely describing the specific

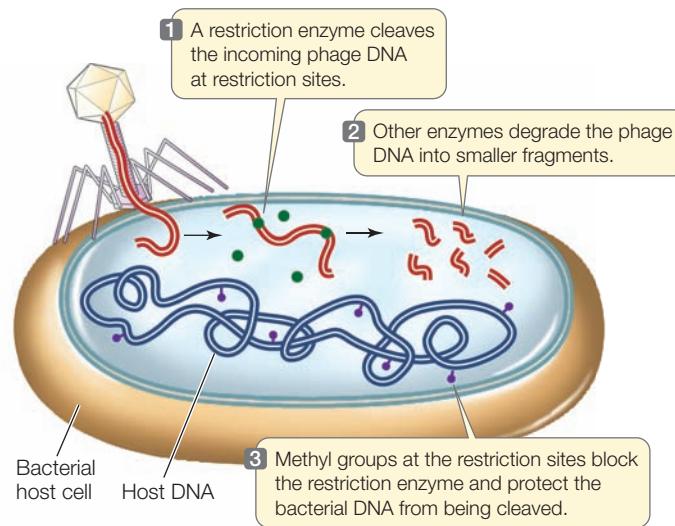
DNA changes that lead to specific protein changes—an area of research called *molecular genetics*. To begin this work, biologists needed tools to analyze DNA molecules for mutations. In this section we will see how some of the numerous naturally occurring enzymes that cleave DNA have now become one of the most important tools used in molecular genetics laboratories.

Restriction enzymes cleave DNA at specific sequences

All organisms, including bacteria, must have ways of dealing with their enemies. As we saw in Section 13.1, bacteria are attacked by viruses called bacteriophage. These viruses inject their genetic material into the host cell and turn it into a virus-producing factory, eventually killing the cell. Some bacteria defend themselves against such invasions by producing **restriction enzymes** (also known as *restriction endonucleases*), which cut double-stranded DNA molecules—such as those injected by bacteriophage—into smaller, noninfectious fragments (Figure 15.7). These enzymes break the bonds of the DNA backbone between the 3' hydroxyl group of one nucleotide and the 5' phosphate group of the next nucleotide. This cutting process is called **restriction digestion**.

There are many such restriction enzymes, each of which cleaves DNA at a specific sequence of bases called a **recognition sequence** or a **restriction site**. Most recognition sequences are 4–6 base pairs long. The sequence is recognized through the principles of protein–DNA interactions (see Section 13.2). That is, the base pairs inside the DNA double helix vary slightly in shape, so that a particular short sequence of base pairs will fit a specific three-dimensional structure on an enzyme.

Why doesn't a restriction enzyme cut the DNA of the bacterial cell that makes it? One way that the cell protects itself is by modifying the restriction sites on its own DNA. Specific modifying enzymes called *methylases* add methyl ($-\text{CH}_3$) groups to certain bases at the restriction sites on the host's DNA after it has been replicated. The methylation of the host's bases makes



15.7 Bacteria Fight Invading Viruses by Making Restriction Enzymes

the recognition sequence unrecognizable to the restriction enzyme. But unmethylated phage DNA is efficiently recognized and cleaved.

Bacterial restriction enzymes can be isolated from the cells that make them and used as biochemical reagents in the laboratory to give information about the nucleotide sequences of DNA molecules from other organisms. If DNA from any organism is incubated in a test tube with a restriction enzyme (along with buffers and salts that help the enzyme to function), that DNA will be cut wherever the restriction site occurs. A specific sequence of bases defines each restriction site. For example, the enzyme *Eco*RI (named after its source, a strain of the bacterium *E. coli*) cuts DNA only where it encounters the following paired sequence in the DNA double helix:



Note that this sequence is palindromic, like the word “mom,” in that the opposite strands have the same sequences when they are read from their 5' ends. The *Eco*RI enzyme has two identical active sites on its two subunits, which cleave the two strands simultaneously between the G and the A of each strand.

The *Eco*RI recognition sequence occurs, on average, about once in every 4,000 base pairs in a typical prokaryotic genome, or about once per four prokaryotic genes. So *Eco*RI can chop a large piece of DNA into smaller pieces containing, on average, just a few genes. Using *Eco*RI in the laboratory to cut small genomes, such as those of viruses that have tens of thousands of base pairs, may result in a few fragments. For a huge eukaryotic chromosome with tens of millions of base pairs, a very large number of fragments will be created.

Of course, “on average” does not mean that the enzyme cuts all stretches of DNA at regular intervals. For example, the *Eco*RI recognition sequence does not occur even once in the 40,000 base pairs of the T7 phage genome—a fact that is crucial to the survival of this virus, since its host is *E. coli*. Fortunately for *E. coli*, the *Eco*RI recognition sequence does appear in the DNA of other bacteriophage.

Hundreds of restriction enzymes (all with unique recognition sequences) have been purified from various microorganisms. In the laboratory, different restriction enzymes can be used to cut samples of DNA from the same source. Thus restriction enzymes can be used to cut a sample of DNA in many different, specific places. The fragments formed can be used to create a physical map of the intact DNA molecule. Before DNA sequencing technology became automated and widely available, this was the principal way that DNA from different organisms was mapped and characterized.

Restriction enzyme digestion is used very widely today to manipulate DNA in the laboratory, and to identify and analyze point mutations. Fragments of DNA from different organisms can be amplified using the polymerase chain reaction (PCR; see Section 13.5). Even between closely related individuals, these amplified fragments often contain variations in DNA sequences (due in most cases to silent mutations). If these variations affect

restriction sites, then digestion of the fragments with restriction enzymes can be used to distinguish between the samples. Restriction enzymes are also used to cut DNA for use in genetic engineering experiments, and in many other types of experiments aimed at understanding how organisms function at the molecular level.

Gel electrophoresis separates DNA fragments

After a laboratory sample of DNA has been cut with a restriction enzyme, the DNA is in fragments, which must be separated to identify (map) where the cuts were made. Because the recognition sequence does not occur at regular intervals, the fragments are not all the same size, and this property provides a way to separate them from one another. Separating the fragments is necessary to determine the number and molecular sizes (in base pairs) of the fragments produced, or to identify and purify an individual fragment for further analysis or for use in an experiment.

A convenient way to separate or purify DNA fragments is by **gel electrophoresis**. Samples containing the fragments are placed in wells at one end of a semisolid gel (usually made of agarose or polyacrylamide polymers) and an electric field is applied to the gel (Figure 15.8). Because of its phosphate groups, DNA is negatively charged at neutral pH; therefore, because opposite charges attract, the DNA fragments move through the gel toward the positive end of the field. Because the spaces between the polymers of the gel are small, small DNA molecules can move through the gel faster than larger ones. Thus, DNA fragments of different sizes separate from one another and can be detected with a dye. This gives us three types of information:

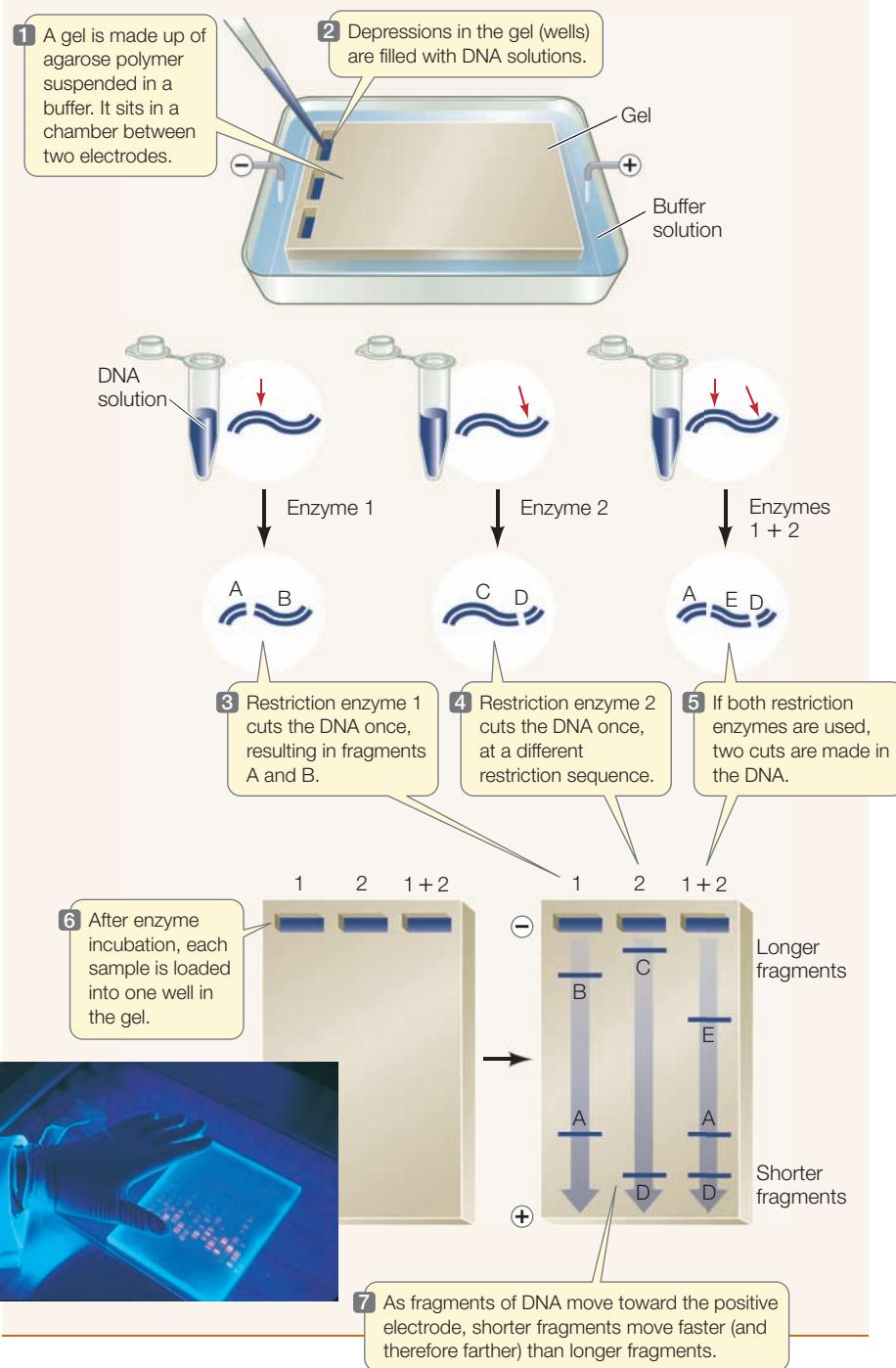
- *The number of fragments.* The number of fragments produced by digestion of a DNA sample with a given restriction enzyme depends on how many times that enzyme's recognition sequence occurs in the sample. Thus gel electrophoresis can provide some information about the presence of specific DNA sequences in the DNA sample.
- *The sizes of the fragments.* DNA fragments of known size are often placed in one well of the gel to provide a standard for comparison. This tells us how large the DNA fragments in the other wells are. By comparing the fragment sizes obtained with two or more restriction enzymes, the locations of their recognition sites relative to one another can be worked out (mapped).
- *The relative abundance of a fragment.* In many experiments, the investigator is interested in how much DNA is present. The relative *intensity* of a band produced by a specific fragment can indicate the amount of that fragment.

After separation on a gel, a fragment with a specific DNA sequence can be revealed with a single-stranded DNA probe (as we will see later in this chapter; see Figure 15.16). The gel region containing the desired fragment (in size or sequence) can be cut out as a lump of gel, and the pure DNA fragment can then be removed from the gel by diffusion into a small volume of water. This fragment can then be analyzed in terms of sequence or amplified and used experimentally.

TOOLS FOR INVESTIGATING LIFE

15.8 Separating Fragments of DNA by Gel Electrophoresis

A mixture of DNA fragments is placed in a gel and an electric field is applied across the gel. The negatively charged DNA moves toward the positive end of the field, with smaller molecules moving faster than larger ones. After minutes to hours for separation, the electric power is shut off and the separated fragments can be analyzed.



yourBioPortal.com

GO TO Animated Tutorial 15.1 • Separating Fragments of DNA by Gel Electrophoresis

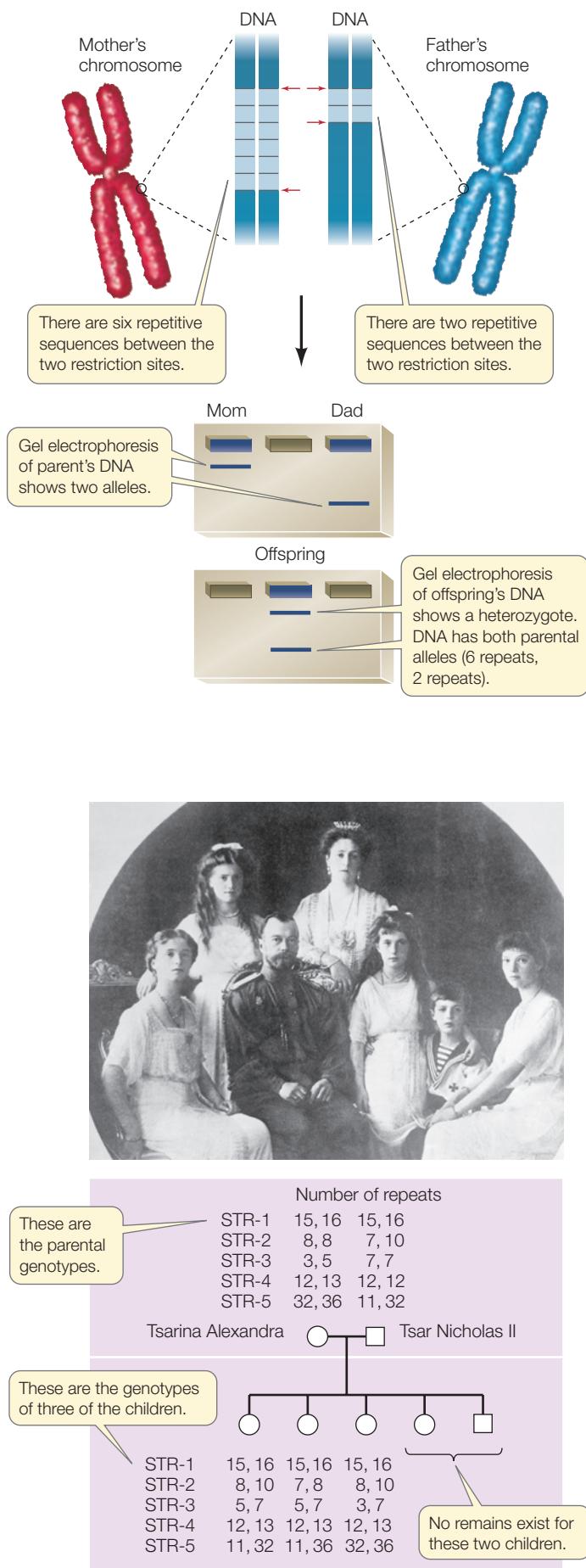
DNA fingerprinting uses restriction analysis and electrophoresis

The two methods we have just described—restriction digestion to cut DNA into fragments and gel electrophoresis to separate them by size—are techniques used in **DNA fingerprinting**, which identifies individuals based on their DNA profiles. DNA fingerprinting works best with sequences that are highly polymorphic—that is, sequences that have multiple alleles (due to many point mutations during the evolution of the organism) and are therefore likely to be different in different individuals. Two types of polymorphisms are especially informative:

- **Single nucleotide polymorphisms (SNPs)**; pronounced “snips”) are inherited variations involving a single nucleotide base (so SNPs are point mutations). These polymorphisms have been mapped for many organisms. If one parent is homozygous for A at a certain point on the genome, and the other parent has a G at that point, the offspring will be heterozygous: one chromosome will have A at that point and the other will have G. If a SNP occurs in a restriction enzyme recognition site, such that one variant is recognized by the enzyme and the other isn’t, then individuals can be distinguished from one another by amplifying a DNA fragment containing that site from a sample of total DNA isolated from each individual. The fragments are then cut with the restriction enzyme and analyzed by gel electrophoresis.

- **Short tandem repeats (STRs)** are short, repetitive DNA sequences that occur side by side on the chromosomes, usually in the noncoding regions. These repeat patterns, which contain 1–5 base pairs, are also inherited. For example, at a particular locus on chromosome 15 there may be an STR of “AGG.” An individual may inherit an allele with six copies of the repeat (AGGAGGAGGAGGAGGAGG) from her mother and an allele with two copies (AGGAGG) from her father. Again, PCR is used to amplify DNA fragments containing these repeat sequences, and the fragments are distinguished by gel electrophoresis (**Figure 15.9**).

The method of DNA fingerprinting used most commonly today involves STR analysis. When several different STR loci, each with numerous alleles, are analyzed, an individual’s unique pat-



15.9 DNA Fingerprinting with Short Tandem Repeats A particular STR locus can be analyzed to determine the number of repeat sequences that were inherited by an individual from each parent. The two alleles can be identified in an electrophoresis gel on the basis of their sizes. When several STR loci are analyzed, the pattern can constitute a definitive identification of an individual.

tern becomes apparent. The Federal Bureau of Investigation in the United States uses 13 STR loci in its Combined DNA Index System (CODIS) database.

DNA fingerprinting can be used in forensics (crime investigation) to help prove the innocence or guilt of a suspect. It has other uses, as well.

A fascinating example demonstrates the use of DNA fingerprinting in the analysis of historical events. Three hundred years of rule by the Romanov dynasty in Russia ended on July 16, 1918, when Tsar Nicholas II, his wife, and their five children were executed by a firing squad during the Communist revolution. A report that the bodies had been burned to ashes was never questioned until 1991, when a shallow grave with several skeletons was discovered several miles from the presumed execution site. DNA fingerprinting of bone fragments found in this grave indicated that they came from an older man, a woman, and three female children, who all were clearly related to one another (Figure 15.10) and were also related to several living descendants of the Tsar. The accuracy and specificity of these methods gave historical and cultural closure to a major event in the twentieth century.

The DNA barcode project aims to identify all organisms on Earth

One of the most exciting aspects of DNA technology for biologists is its potential to identify species, varieties, and even individual organisms from their DNA. In order to repeat experiments and report scientific results, it is essential that biologists know exactly what species or varieties they are studying. However, different organisms can sometimes look very much alike in nature. About 1.7 million species have been named and described, but about ten times that number probably have yet to be identified. A proposal to use DNA technology to identify known species and detect the unknown ones has been endorsed by a large group of scientific organizations known as the Consortium for the Barcode of Life (CBOL).

Evolutionary biologist Paul Hebert at the University of Guelph in Ontario, Canada was walking down the aisle of a supermarket in 1998 when he noticed the barcodes on all the packaged foods. This gave him an idea to identify each species with a “DNA barcode” that is based on a short sequence from a sin-

15.10 DNA Fingerprinting of the Russian Royal Family The skeletal remains of Tsar Nicholas II, his wife Alexandra, and three of their children were found in 1991 and subjected to DNA fingerprinting. Five STRs were tested. The results can be interpreted by looking at the inheritance of alleles from each parent in the children. In STR-2, for example, the parents had genotypes 8,8 (homozygous) and 7,10 (heterozygous). The three children all inherited type 8 from Alexandra and either type 7 or type 10 from Nicholas.

15.11 A DNA Barcode A 650- to 750-base-pair region of the cytochrome oxidase gene can be amplified by PCR from any organism and then sequenced. This knowledge is used to make a bar code in which each of the four DNA bases is represented as a different color. Such a species barcode permits accurate and rapid identification of a particular species for experimental, ecological or evolutionary studies.

gle gene. The gene he chose is the cytochrome oxidase gene, a component of the respiratory chain that is present in most organisms. Because this gene mutates readily, there are many allelic differences between species. A fragment of 650–750 base pairs in this gene is being sequenced for all organisms, and so far sufficient variation has been detected to make it diagnostic for each species (**Figure 15.11**).

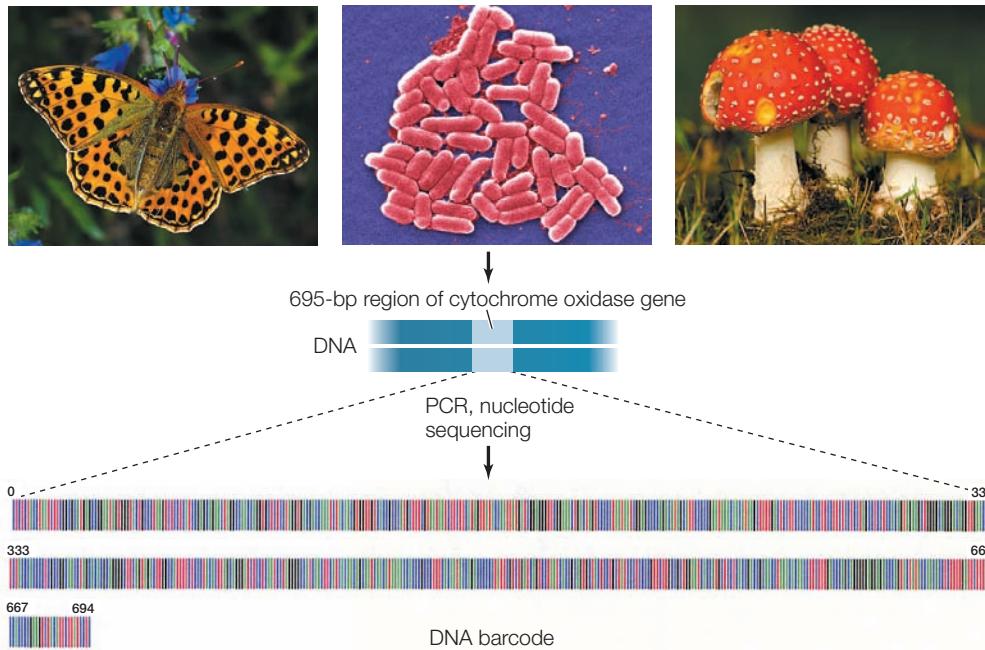
Once the DNA of the targeted gene fragment has been sequenced for all known species, a simple device for conducting field analyses can be developed. The barcode project has the potential to advance biological research on evolution, to track species diversity in ecologically significant areas, to help identify new species, and even to detect undesirable microbes or bioterrorism agents.

15.2 RECAP

Large DNA molecules can be cut into smaller pieces by restriction digestion and then sorted by gel electrophoresis. PCR is used to amplify sequences of interest from complex samples. These techniques are used in DNA fingerprinting to analyze DNA polymorphisms for the purpose of identifying individuals. Scientists hope to identify all species using DNA analysis.

- How does a restriction enzyme recognize a restriction site on DNA? **See p. 323**
- How does gel electrophoresis separate DNA fragments? **See p. 324 and Figure 15.8**
- What are STRs and how are they used to identify individuals? **See pp. 325–326**

We have seen that molecular methods can be used to identify individuals because of mutations in their DNA. Many of the STRs and SNPs used in these analyses do not occur in protein-coding regions, and so probably do not affect the phenotype. Nevertheless, they are mutations— inherited changes in the DNA. We now turn to mutations that affect phenotype, using humans as our model organism.



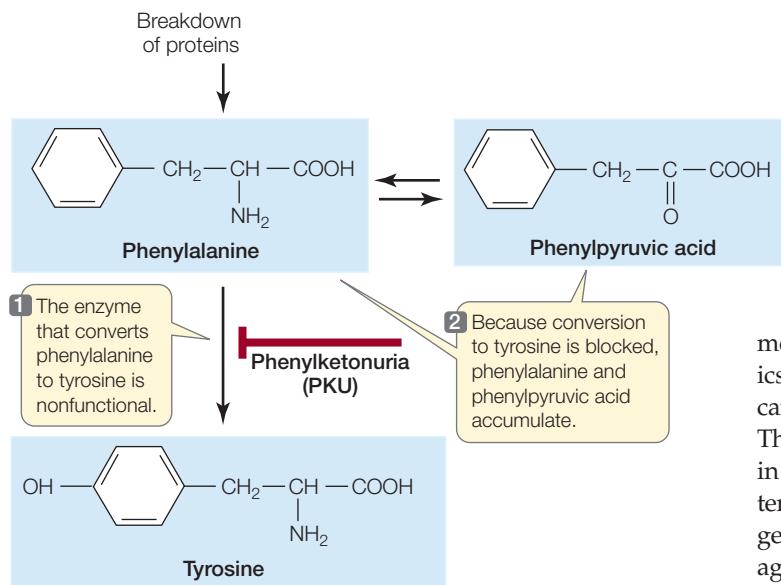
15.3 How Do Defective Proteins Lead to Diseases?

The biochemistry that relates genotype (DNA) and phenotype (proteins) has been most completely described for model organisms, such as the prokaryote *E. coli* and the eukaryotes yeast and *Drosophila*. While the details vary, there is great similarity in the fundamental processes among these forms of life. These similarities have permitted the application of knowledge and methods discovered using these model organisms to the study of human biochemical genetics. Of particular interest are the effects of mutations on human phenotypes, sometimes leading to diseases.

Genetic mutations may make proteins dysfunctional

Genetic mutations are often expressed phenotypically as proteins that differ from normal (wild-type) proteins. Abnormalities in enzymes, receptor proteins, transport proteins, structural proteins, and most of the other functional classes of proteins have all been implicated in genetic diseases.

DYSFUNCTIONAL ENZYMES In 1934, the urine of two mentally retarded young siblings was found to contain phenylpyruvic acid, an unusual by-product of the metabolism of the amino acid phenylalanine. It was not until two decades later, however, that the complex clinical phenotype of the disease that afflicted these children, called *phenylketonuria* (PKU), was traced back to its molecular cause. The disease resulted from an abnormality in a single enzyme, phenylalanine hydroxylase (**Figure 15.12**). This enzyme normally catalyzes the conversion of dietary phenylalanine to tyrosine, but it was not active in the livers of PKU patients. Lack of this conversion led to excess phenylalanine in the blood and explained the accumulation of phenylpyruvic acid. Later, the amino acid sequence of phenylalanine hydroxylase (PAH) in normal people was compared with the amino acid sequences



from individuals with PKU. Many people with PKU had tryptophan instead of arginine at position 408 of this long polypeptide chain of 452 amino acids (Table 15.1).

The exact cause of mental retardation in PKU remains elusive, although, as we will see later in this chapter, it can be prevented. We can, however, understand why most people with PKU have light skin and hair color. The pigment melanin, which is responsible for dark skin and hair, is made from tyrosine, which people with PKU cannot synthesize adequately.

Hundreds of human genetic diseases that result from enzyme abnormalities have been discovered, many of which lead to mental retardation and premature death. Most of these diseases are rare; PKU, for example, shows up in one out of every 12,000 newborns. But these diseases are just the tip of the mutation iceberg. Some mutations result in amino acid changes that have no obvious clinical effects. In fact, amino acid differences among individuals have been detected in at least 30 percent of all human proteins whose sequences are known. Thus polymorphism does not necessarily mean disease. There can be numerous alleles of a gene, some producing proteins that function normally, while others produce variants that cause disease—as we will now see for hemoglobin.

15.12 One Gene, One Enzyme Phenylketonuria is caused by an abnormality in a specific enzyme that metabolizes the amino acid phenylalanine. Knowing the molecular causes of such single-gene, single-enzyme metabolic diseases can aid researchers in developing screening tests as well as treatments.

ABNORMAL HEMOGLOBIN The first human genetic disease known to be caused by an amino acid sequence abnormality was sickle-cell disease. This blood disorder most often afflicts people whose ancestors came from the tropics or from the Mediterranean. About 1 in 655 African-Americans are homozygous for the sickle allele and have the disease. The abnormal allele produces abnormal hemoglobin that results in sickle-shaped red blood cells (see Figure 15.3). These cells tend to block narrow blood capillaries, especially when the oxygen concentration of the blood is low. The result is tissue damage and eventually death by organ failure.

Recall that human hemoglobin is a protein with quaternary structure, containing four globin subunits—two α chains and two β chains—as well as the pigment heme (see Figure 3.10). In sickle-cell disease, one of the 146 amino acids in the β -globin chain is abnormal: at position 6, the normal glutamic acid has been replaced by valine. This replacement changes the charge of the protein (glutamic acid is negatively charged and valine is neutral), causing it to form long, needle-like aggregates in the red blood cells. The phenotypic result is anemia, a deficiency of normal red blood cells and an impaired ability of the blood to carry oxygen.

Because hemoglobin is easy to isolate and study, its variations in the human population have been extensively documented (Figure 15.13). Hundreds of single amino acid alterations in β -globin have been reported. For example, at the same position that is mutated in sickle-cell disease (resulting in hemoglobin S), the normal glutamic acid may be replaced by lysine, causing hemoglobin C disease. In this case, the resulting anemia is usually not severe. Many alterations of hemoglobin do not affect the protein's function. That is fortunate, because about 5 percent of all humans are carriers for one of these variants.

There are hundreds of inherited diseases in humans in which the primary phenotypes are caused by specific mutations leading to protein abnormalities. Some of the more common examples are listed in Table 15.2. These mutations can be domi-

TABLE 15.1
Two Common Mutations That Cause Phenylketonuria

	NORMAL CODON 408	MUTANT CODON 408 (20% OF PKU CASES)	NORMAL CODON 280	MUTANT CODON 280 (2% OF PKU CASES)
Length of PAH protein	452 amino acids	452 amino acids	452 amino acids	452 amino acids
DNA at codon	xxCGGxx xxGCCxx	xxTGGxx xxACCxx	xxGAAxx xxCTTxx	xxAAAx xxTTTxx
mRNA at codon	xxCGGxx	xxUGGxx	xxGAAxx	xxAAAxx
Amino acid at codon	Arginine	Tryptophan	Glutamic acid	Lysine
Active PAH enzyme?	Yes	No	Yes	No

TABLE 15.2
Some Human Genetic Diseases

DISEASE NAME	INHERITANCE PATTERN; FREQUENCY	GENE MUTATED; PROTEIN PRODUCT	CLINICAL PHENOTYPE
Familial hypercholesterolemia	Autosomal codominant; 1 in 500 heterozygous	<i>LDLR</i> ; low-density lipoprotein receptor	High blood cholesterol, heart disease
Cystic fibrosis	Autosomal recessive; 1 in 4000	<i>CFTR</i> ; chloride ion channel in membrane	Immune, digestive, and respiratory illness
Duchenne muscular dystrophy	Sex-linked recessive; 1 in 3500 males	<i>DMD</i> ; the muscle membrane protein dystrophin	Muscle weakness
Hemophilia A	Sex-linked recessive; 1 in 5000 males	<i>HEMA</i> ; factor VIII blood clotting protein	Inability to clot blood after injury, hemorrhage

nant, codominant, or recessive, and some are sex-linked. Before we examine how these diseases can be analyzed at the molecular level, we turn briefly to a fascinating exception to the association between genes and proteins.

Prion diseases are disorders of protein conformation

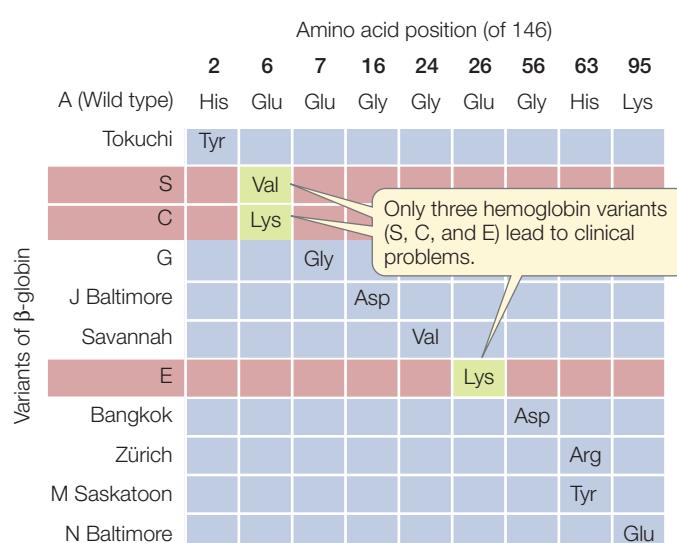
Transmissible spongiform encephalopathies (TSEs) are degenerative brain diseases that occur in many mammals, including humans. The brain gradually develops holes, making it look like a sponge. Scrapie is a TSE that has been known for 250 years. It causes affected sheep and goats to show the abnormal behavior of rubbing ("scraping") the wool off their bodies (as well as causing more severe neurological problems). In the 1980s, a TSE that appeared in cows in Britain was traced to the cows having eaten products from sheep that had scrapie. These cows would shake and rub their bodies against fences, and their staggering led farmers to dub them "mad cows." In the 1990s, some people who ate beef from these cows got a human version of the disease, dubbed "mad cow disease" by the media. Those with the disease eventually died.

At first, viruses were suspected to cause TSEs. But when Tikva Alper at Hammersmith Hospital, London, treated infectious extracts with high doses of ultraviolet light to inactivate nucleic acids, they still caused TSEs. She proposed that the causative agent was a protein, not a virus. Later, Stanley Prusiner at the University of California purified the protein responsible and showed it to be free of DNA or RNA. He called it a *proteinaceous infective particle*, or **prion**. This is a violation of the central dogma of molecular biology (DNA → RNA → protein; see Chapter 14), because in this case the protein was "doing it all." There was no genetic material involved.

This is a rare case of a mutant phenotype without a mutant gene. Normal brain cells contain a membrane protein called PrP^c. A protein with the *same amino acid sequence* is present in TSE-affected brain tissues, but that protein, called PrP^{sc}, has a different three-dimensional shape (Figure 15.14). Thus TSEs are not caused by a mutated gene (the primary structures of the two proteins are the same), but are somehow caused by an alteration in protein conformation. The altered three-dimensional structure of the protein has profound effects on its function in the cell. PrP^{sc} is insoluble, and it piles up as fibers in brain tissue, causing cell death.

How can the exposure of a normal cell to material containing PrP^{sc} result in a TSE? The abnormal PrP^{sc} protein seems to induce a conformational change in the normal PrP^c protein so that it, too, becomes abnormal. Just how this conversion occurs, and how it causes a TSE, is unclear.

To try to understand how TSEs develop, scientists are asking "What is the *normal* role of the prion protein?" Recently, it was shown that in the brain the prion protein blocks a key enzyme in the synthesis of a protein called β -amyloid. This is the protein that accumulates in the brains of patients with Alzheimer's disease. People with early-onset Alzheimer's (age 40) have less PrP^c in their brains than people who age normally. So the PrP^c pro-



15.13 Hemoglobin Polymorphism Each of these mutant alleles codes for a protein with a single amino acid change in the 146-amino acid chain of β -globin. Only three of the hundreds of known variants of β -globin, shown on the left, are known to lead to clinical abnormalities. "S" is the sickle-cell anemia allele.

15.14 Prion Diseases are Disorders of Protein Conformation A normal membrane protein in brain cells (PrP^{C} , left) can be converted to the disease-causing form (PrP^{Sc} , right), which has a different three-dimensional structure.

tein appears to play a role in protecting against Alzheimer's disease. Other functions for this protein are also being discovered, but the mechanism by which it appears to spread TSE disease is not yet understood.

Prions are an unusual phenomenon in human disease. The vast majority of inherited diseases are caused by mutations in genes that reduce the levels of their protein products, or make the proteins dysfunctional. But the expression of these genes, like that of all genes, is influenced by the environment.

Most diseases are caused by multiple genes and environment

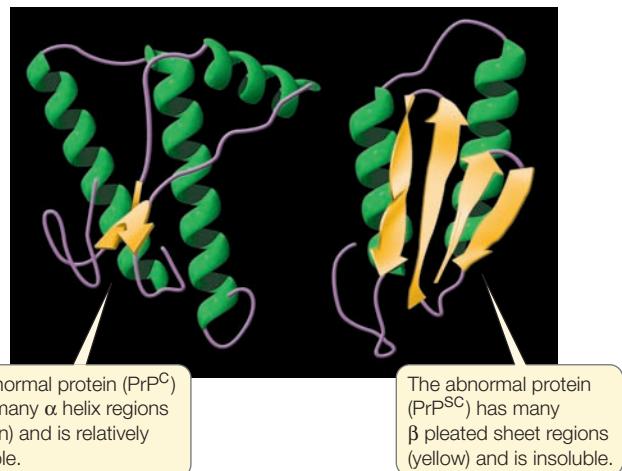
The human diseases for which clinical phenotypes can be traced to a single altered protein and its altered gene may number in the thousands. Taken together, these diseases have a frequency of about one percent in the human population.

Far more common, however, are diseases that are **multifactorial**; that is, diseases that are caused by the interactions of many genes and proteins with one or more factors in the environment. When studying genetics, we tend to call individuals either normal (wild type) or abnormal (mutant); however, in reality every individual contains thousands or millions of genetic variations that arose through mutations. Our susceptibility to disease is often determined by complex interactions between these genotypes and factors in the environment, such as the foods we eat or the pathogens we encounter. For example, a complex set of genotypes determine who among us can eat a high-fat diet and not experience a heart attack, or will succumb to disease when exposed to infectious bacteria. Estimates suggest that up to 60 percent of all people are affected by diseases that are *genetically influenced*. Identifying these genetic influences is a major task of molecular medicine and human genome sequencing.

15.3 RECAP

Many genetic mutations are expressed as nonfunctional enzymes, structural proteins, or membrane proteins. Human genetic diseases may be inherited in dominant, codominant, or recessive patterns, and they may be sex-linked.

- Describe an example of an abnormal protein in humans that results from a genetic mutation and causes a disease. **See pp. 327–328**
- Describe an example of an abnormal protein in humans that results from a genetic mutation and does not cause a disease. **See p. 328**
- How is the brain cell membrane protein PrP^{C} related to diseases caused by prions? **See p. 329**



The abnormal proteins that cause disease result (with the exception of TSEs) from genetic mutations. We now turn to the identification of such mutations, an important task for molecular medicine.

15.4 What DNA Changes Lead to Genetic Diseases?

We have seen for diseases such as PKU and sickle-cell anemia that the clinical phenotype of inherited diseases could be traced to individual proteins, and that the genes could then be identified. With the advent of new ways to identify DNA variations, a new pattern of human genetic analysis has emerged. In these cases, the clinical phenotype is first related to a DNA variation, and then the protein involved is identified. This pattern of discovery is called **reverse genetics**, because it proceeds in the opposite direction to genetic analyses done before the mid-1980s. For example, in sickle-cell anemia, the protein abnormality in hemoglobin was described first (a single amino acid change), and then the gene for β -globin was isolated and the DNA mutation was pinpointed.

clinical phenotype → protein phenotype → gene

On the other hand, for cystic fibrosis (see Table 15.2), a mutant version of the gene *CFTR* was isolated first, and then the protein was characterized:

clinical phenotype → gene → protein phenotype

Whichever approach is used, final identification of the protein(s) involved in a disease is important in designing specific therapies.

Genetic markers can point the way to important genes

To identify a mutant gene by reverse genetics, close linkage to a marker sequence is used. To understand this linkage, imagine an astronaut looking down from space, trying to find her son on a park bench on Chicago's North Shore. The astronaut first picks out reference points—landmarks that will lead her to the park. She recognizes the shape of North America, then moves to Lake Michigan, then the Willis Tower, and so on. Once she

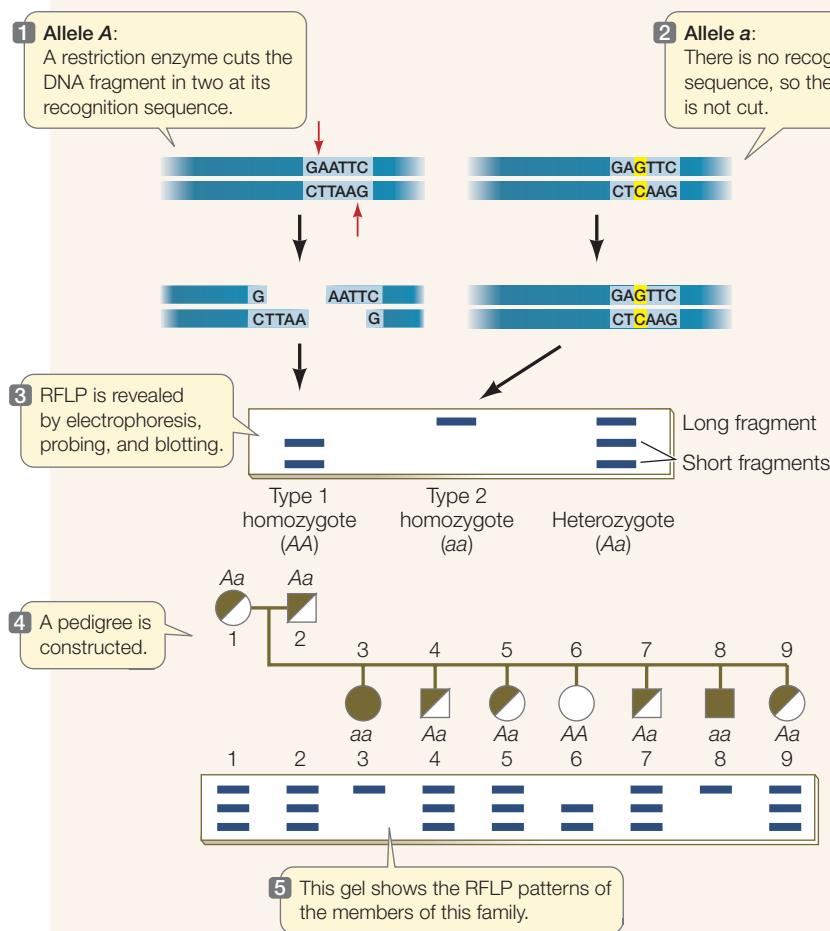
has zeroed in on the North Shore Park, she can use advanced optical instruments to find her son. The reference points for gene isolation are the **genetic markers**.

- Knowledge of at least *two mutations* is needed. One mutation determines the disease phenotype and the other mutation is a closely linked “marker mutation” that does not affect the disease phenotype but is easy to identify. In early genetic studies, markers that produced visible phenotypes were used to follow the inheritance of important traits. Today, single nucleotide polymorphisms (SNPs) or STRs are usually used.
- *Genetic linkage* is the co-inheritance of the marker and the disease-causing allele. If they are always together, they must be close together on the chromosome.

TOOLS FOR INVESTIGATING LIFE

15.15 RFLP Markers

Restriction fragment length polymorphisms are differences in DNA sequences that serve as genetic markers. Linkage studies can be used to isolate genes involved with diseases.



A key requirement for a genetic marker is that it has allelic polymorphisms (differences in sequence) that are identifiable by current methods of rapid DNA analysis. As we saw in Section 15.2, an STR can have varying numbers of a short repeat sequence, and thus there are multiple alleles of these markers. We also saw in Section 15.2 that restriction enzymes can be used to identify SNPs, provided the SNP occurs within a restriction site. Restriction enzymes can also be used to identify mutations such as insertions or deletions if the affected sequences contain restriction sites. We will examine in more detail the use of restriction enzymes to identify genetic polymorphisms, and other SNP markers, before returning to the discussion of human genes and their abnormalities.

RESTRICTION FRAGMENT LENGTH POLYMORPHISMS

As Section 15.2 describes, restriction enzymes cut DNA molecules at specific recognition sequences. On a particular human chromosome, a given restriction enzyme may make thousands of cuts, producing many DNA fragments. The enzyme *Eco*RI, for example, cuts DNA at



Suppose this recognition sequence exists in a certain stretch of human chromosome 7. The restriction enzyme will cut this stretch once and make two fragments of DNA. Now suppose that, in some people, this sequence contains a SNP and is mutated as follows:



This sequence will not be recognized by the restriction enzyme; thus it will remain intact and yield one larger fragment of DNA.

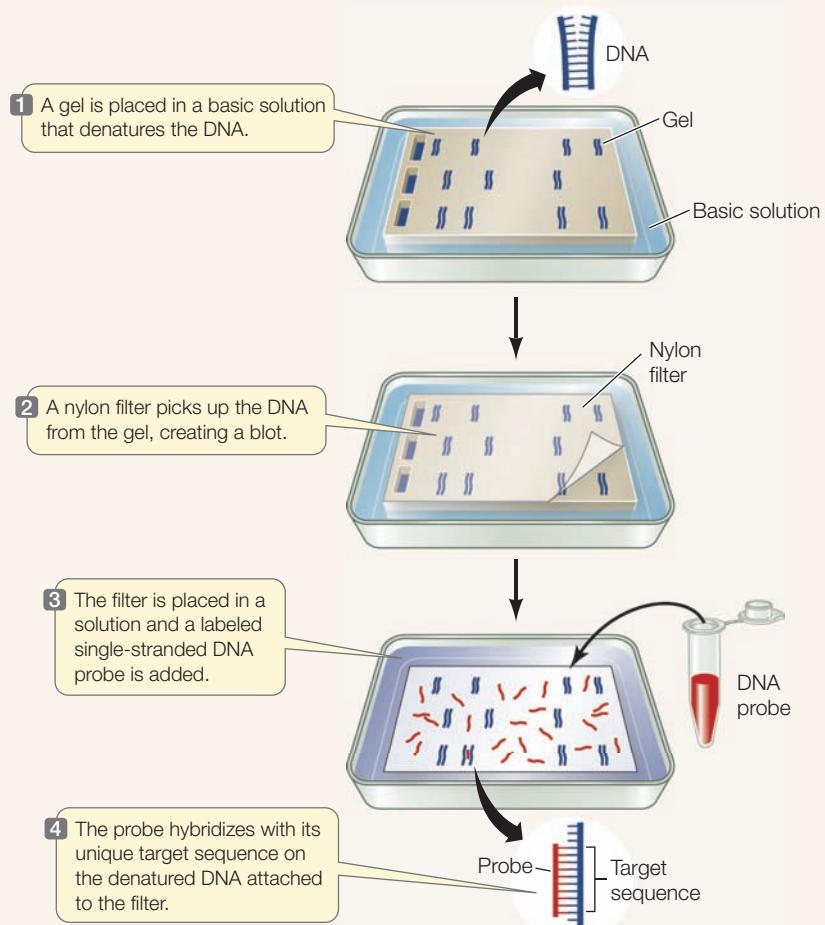
Differences in DNA sequences due to mutations in restriction sites are called **restriction fragment length polymorphisms (RFLPs)** (Figure 15.15). They can be easily visualized as bands on an electrophoresis gel. An RFLP band pattern is inherited in a Mendelian fashion and can be followed through a pedigree. Thousands of such markers have been described for humans and many other organisms.

Before the advent of PCR technology, the only way to analyze RFLPs was by digesting total genomic DNA samples with restriction enzymes. These samples contain thousands of DNA fragments of various sizes. In order to visualize a particular fragment, the DNA from the gel is transferred (blotted) onto a nylon membrane, denatured to separate the double-stranded molecules, and mixed with a single-stranded DNA fragment (*a probe*) containing at least part of the sequence within the RFLP fragment of interest (Figure 15.16). The probe hybridizes (by base pairing) with the DNA band containing the RFLP. Because the probe is “la-

TOOLS FOR INVESTIGATING LIFE

15.16 Analyzing DNA Fragments by DNA Gel Blotting

A probe can be used to locate a specific DNA fragment on an electrophoresis gel.



beled" with a radioactive isotope or a chemical tag, the DNA fragment containing the RFLP can be seen among the thousands of other fragments on the blot. This technology was used to create "RFLP maps" of the human genome and of genomes of many other organisms.

SINGLE NUCLEOTIDE POLYMORPHISMS As noted in Section 15.2, single nucleotide polymorphisms (SNPs) are widespread in eukaryotic genomes. There is roughly one SNP for every 1,330 base pairs in the human genome. Not all SNPs occur within restriction sites, but those that don't can still be used as markers. SNPs can be detected by direct sequence comparisons, or by PCR amplification using primers that contain one version of the SNP, so that only one allele will be amplified efficiently. SNPs can also be detected using sophisticated chemical methods such as mass spectrometry (see Section 17.5).

Genetic markers such as STRs, RFLPs, and SNPs can be used as landmarks to find genes of interest if the genes also have alleles and are polymorphic. The key to this method is the well-

known observation that if two genes are located near each other on the same chromosome, they are usually passed on together from parent to offspring. The same holds true for any pair of genetic markers. The idea is to find markers that are close by on the chromosome to get to the gene of interest.

To narrow down the location of a gene, a scientist must find a marker and a gene that are *always inherited together*. To do this, family medical histories are taken and pedigrees are constructed. If a genetic marker and a genetic disease are inherited together, then they must be near each other on the same chromosome. Unfortunately, "near each other" still might be as much as several million base pairs apart. The process of locating a gene is thus similar to that of the astronaut looking for her son: the first landmarks lead to only an approximate location.

How are markers identified that are more closely linked to the gene of interest? Now that the human genome has been fully sequenced (see Chapter 17), the task is much easier than it was just a decade ago. Sequence information from the chromosomal region near the linked marker is used to develop additional SNP-based or STR markers. These are tested to identify markers that are tightly linked to the disease phenotype. Eventually, the region of DNA containing the gene can be narrowed down to a few hundred thousand base pairs.

Once a linked DNA region is identified, many methods are available to identify the actual gene responsible for a genetic disease. The complete sequence of the region can be searched for candidate genes, using information available from databases of genome sequences. With luck a scientist can make an educated guess, based

on biochemical or physiological information about the disease, along with information about the functions of candidate genes, as to which gene is responsible for the disease. The identification of DNA polymorphisms within candidate genes, between diseased and healthy individuals, can also help to narrow down the search. A variety of techniques, such as analyzing mRNA levels of candidate genes in diseased and healthy individuals, are used to confirm that the correct gene has been identified.

The isolation of genes responsible for genetic diseases has led to spectacular advances in the understanding of human biology. For example, the gene responsible for cystic fibrosis was first identified by its close association with a SNP marker. After the gene sequences of people with the disease and people without the disease were compared, a mutation was identified in most patients. This provided a way to test for the presence of that mutation in people, by extracting DNA from cells or tissues that could be easily sampled. Moreover, knowing the gene sequence led to the identification of the protein it codes for and a characterization of the abnormal protein. Treatments that

specifically target this protein are now being devised. Research on the protein in normal people has led to an understanding of its role in the body. So reverse genetics can lead to diagnosis, treatment, and biological understanding.

Disease-causing mutations may involve any number of base pairs

Disease-causing mutations may involve a single base pair (as we saw in the case of hemophilia), a long stretch of DNA (as in cases of Duchenne muscular dystrophy, which we will discuss shortly), multiple segments of DNA (as in fragile-X syndrome), or even entire chromosomes (as we saw with Down syndrome in Section 11.5).

POINT MUTATIONS There are many examples of point mutations in human genetic diseases. In some cases, all of the people with the disease have the same genetic mutation. This is the case with sickle-cell anemia, where a single base pair change in the β -globin gene causes a single amino acid change, which leads to the abnormal protein and phenotype. This is not the situation with most other genetic diseases. For example, over 500 different mutations in the *PAH* (phenylalanine hydroxylase) gene have been discovered in different patients with phenylketonuria (PKU; see Table 15.1). This makes sense if you think about the three-dimensional structure of an enzyme protein and the many amino acid changes that could affect its activity.

LARGE DELETIONS Larger mutations may involve many base pairs of DNA. For example, deletions in the X chromosome that include the gene for the protein dystrophin result in Duchenne muscular dystrophy. Dystrophin is important in organizing the structure of muscles, and people who have only the abnormal form have severe muscle weakness. Sometimes only part of the dystrophin gene is missing, leading to an incomplete but partly functional protein and a mild form of the disease. In other cases, however, deletions span the entire sequence of the gene, so that the protein is missing entirely, resulting in a severe form of the disease. In yet other cases, deletions involve millions of base pairs and cover not only the dystrophin gene but adjacent genes as well; the result may be several diseases simultaneously.

CHROMOSOMAL ABNORMALITIES Chromosomal abnormalities also cause human diseases. Such abnormalities include the gain

or loss of one or more chromosomes (aneuploidy) (see Figure 11.21), loss of a piece of chromosome (deletion), and the transfer of a piece of one chromosome to another chromosome (translocation) (see Figure 15.4). About one newborn in 200 has a chromosomal abnormality. While some of these abnormalities are inherited as preexisting aberrations from one or both parents, others are the result of meiotic events, such as nondisjunction, that occurred during the formation of gametes in one of the parents.

One common cause of mental retardation is *fragile-X syndrome* (Figure 15.17). About one male in 1,500 and one female in 2,000 are affected. These people have a constriction near the tip of the X chromosome. Although the basic pattern of inheritance is that of an X-linked recessive trait, there are departures from this pattern. Not all people with the fragile-X chromosomal abnormality are mentally retarded, as we will see.

Expanding triplet repeats demonstrate the fragility of some human genes

About one-fifth of all males that have the fragile-X chromosomal abnormality are phenotypically normal, as are most of their daughters. But many of those daughters' sons are mentally retarded. In a family in which the fragile-X syndrome appears, later generations tend to show earlier onset and more severe symptoms of the disease. It is almost as if the abnormal allele itself is changing—and getting worse. And that's exactly what is happening.

The gene responsible for fragile-X syndrome (*FMR1*) contains a repeated triplet, CGG, at a certain point in the promoter region. In normal people, this triplet is repeated 6 to 54 times (the average is 29). In mentally retarded people with fragile-X syndrome, the CGG sequence is repeated 200 to 2,000 times.

Males carrying a moderate number of repeats (55–200) show no symptoms and are called premutated. These repeats become more numerous as the daughters of these men pass the chromosome on to their children (Figure 15.18). With more than 200 repeats, increased methylation of the cytosines in the CGG triplets is likely, accompanied by transcriptional inactivation of the *FMR1* gene. The normal role of the protein product of this gene is to bind to mRNAs involved in neuron function and regulate their translation at the ribosome. When the *FMR1* protein is not made in adequate amounts, these mRNAs are not properly translated, and nerve cells die. Their loss often results in mental retardation.

This phenomenon of **expanding triplet repeats** has been found in over a dozen other diseases, such as myotonic dystrophy (involving repeated CTG triplets) and Huntington's disease (in which CAG is repeated). Such repeats, which may be found within a protein-coding region or outside it, appear to be present in many other genes without causing harm. How the repeats expand is not known; one theory is that DNA polymerase may slip after copying a repeat and then fall back to copy it again.

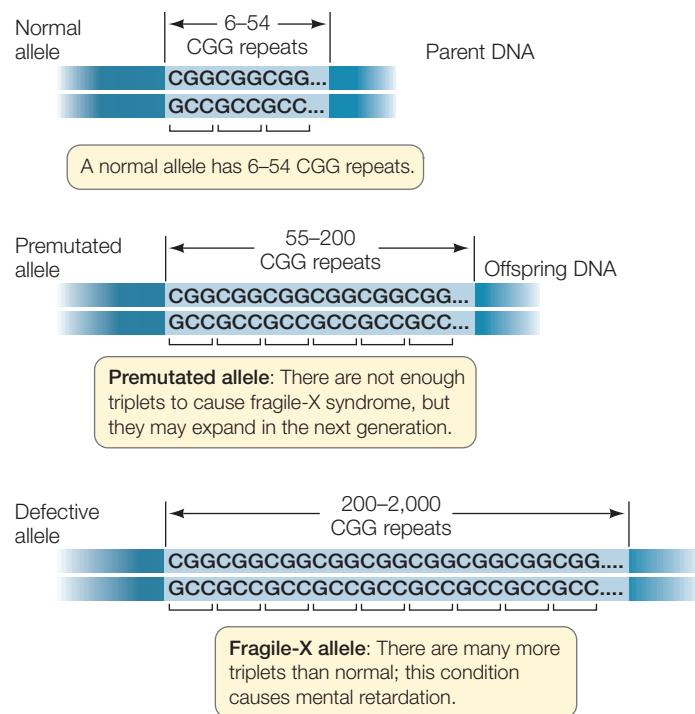


The constriction at the lower tip of this chromosome is the location of the fragile-X abnormality.

15.17 A Fragile-X Chromosome at Metaphase The chromosomal abnormality associated with fragile-X syndrome shows up under the microscope as a constriction in the chromosome. This occurs during preparation of the chromosome for microscopy.

15.18 The CGG Repeats in the FMR1 Gene Expand with Each Generation

Genetic Defect The genetic defect in fragile-X syndrome is caused by 200 or more repeats of the CGG triplet.



15.4 RECAP

Genes involved in disease can be identified by first detecting the abnormal DNA sequence and then the protein that the wild-type allele encodes. Unusual features such as expanding triplet repeats have been detected in the human genome.

- How can a gene be identified before its protein product is known? **See pp. 330–331 and Figure 15.15**
- How do expanding repeats cause genetic diseases? **See p. 333 and Fig. 15.18**

The determination of the precise molecular phenotypes and genotypes of various human genetic diseases has made it possible to diagnose these diseases even before symptoms first appear. Let's take a detailed look at some of these genetic screening techniques.

15.5 How Is Genetic Screening Used to Detect Diseases?

Genetic screening is the use of a test to identify people who have, are predisposed to, or are carriers of a genetic disease. It can be done at many times of life and used for many purposes.

- Prenatal screening can be used to identify an embryo or fetus with a disease so that medical intervention can be applied or decisions can be made about whether or not to continue the pregnancy.

- Newborn babies can be screened so that proper medical intervention can be initiated quickly for those babies who need it.

- Asymptomatic people who have a relative with a genetic disease can be screened to determine whether they are carriers of the disease or are likely to develop the disease themselves.

Genetic screening can be done at the level of either the phenotype or the genotype.

Screening for disease phenotypes involves analysis of proteins

At the level of the phenotype, genetic screening involves examining a protein relevant to the phenotype for abnormal structure or function. Since many proteins are enzymes, low enzyme activity is strongly suggestive of a mutation, as we saw in Section 15.1. Perhaps the best example of this kind of protein screening is a test for phenylketonuria (PKU), which has made it possible to identify the disease in newborns, so that treatment of the disease can be started. It is very likely that you were screened for PKU.

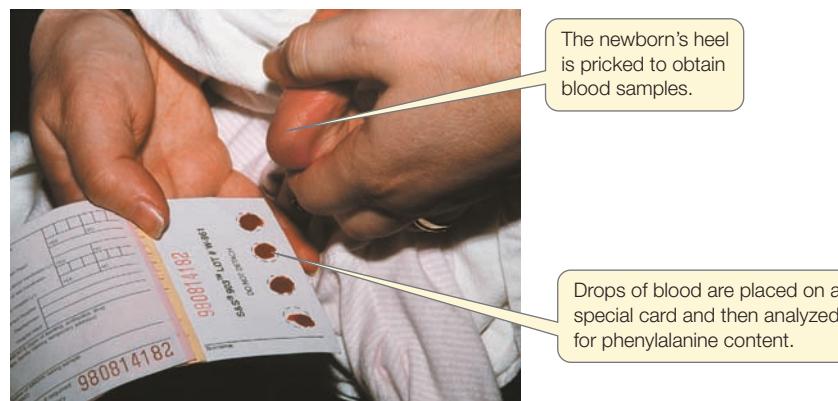
Initially, babies born with PKU have a normal phenotype because excess phenylalanine in their blood before birth diffuses across the placenta to the mother's circulatory system. Since the mother is almost always heterozygous, and therefore has adequate phenylalanine hydroxylase activity, her body metabolizes the excess phenylalanine from the fetus. After birth, however, the baby begins to consume protein-rich food (milk) and to break down some of his or her own proteins. Phenylalanine enters the baby's blood and accumulates. After a few days, the phenylalanine level in the baby's blood may be ten times higher than normal. Within days, the developing brain is damaged, and untreated children with PKU become severely mentally retarded. If detected early, PKU can be treated with a special diet low in phenylalanine to avoid the brain damage that would otherwise result. Thus, early detection is imperative.

Newborn screening for PKU and other diseases began in 1963 with the development of a simple, rapid test for the presence of excess phenylalanine in blood serum (**Figure 15.19**). This method uses dried blood spots from newborn babies and can be automated so that a screening laboratory can process many samples in a day.

Screening using newborn babies' blood is now done for up to 25 genetic diseases. Some are rare, such as maple syrup urine disease, which occurs once in 185,000 births. This disease is caused by a defect in an enzyme that metabolizes certain amino acids, and results in sweet-smelling urine and severe brain damage. Other genetic diseases are more common, such as congenital hypothyroidism, which occurs about once in 4,000 births, and causes reduced growth and mental retardation due to low levels of thyroid hormone. With early intervention, many of these infants can be successfully treated. So it is not surprising that newborn screening is legally mandatory in many countries, including the United States and Canada.

15.19 Genetic Screening of Newborns for Phenylketonuria

A blood test is used to screen newborns for phenylketonuria. Small samples of blood are taken from a newborn's heel. The samples are placed in a machine that measures the phenylalanine concentration in the blood. Early detection means that the symptoms of the condition can be prevented by putting the baby on a therapeutic diet.



DNA testing is the most accurate way to detect abnormal genes

The level of phenylalanine in the blood is an indirect measure of phenylalanine hydroxylase activity in the liver. But how can we screen for genetic diseases that are not detectable by blood tests? What if blood is difficult to obtain, as it is in a fetus? How are genetic abnormalities in heterozygotes, who express the normal protein at some level, identified?

DNA testing is the direct analysis of DNA for a mutation, and it offers the most direct and accurate way of detecting an abnormal allele. Now that the mutations responsible for many human diseases have been identified, any cell in the body can be examined at any time of life for mutations. With the amplification power of PCR, only one or a few cells are needed for testing. These methods work best for diseases caused by only one or a few different mutations.

Consider, for example, two parents who are both heterozygous for the cystic fibrosis allele, who have had a child with the disease, and want a normal child. If treated with the appropriate hormones, the mother can be induced to "superovulate," releasing several eggs. An egg can be injected with a single sperm from her husband and the resulting zygote allowed to divide to the 8-cell stage. If one of these embryonic cells is removed, it can be tested for the presence of the cystic fibrosis allele. If the test is negative, the remaining 7-cell embryo can be implanted in the mother's womb where with luck, it will develop normally.

Such *preimplantation screening* is performed only rarely. More typical are analyses of fetal cells after normal fertilization and implantation in the womb. Fetal cells can be analyzed at about the tenth week of pregnancy by chorionic villus sampling, or during the thirteenth to seventeenth weeks by amniocentesis. In either case, only a few fetal cells are necessary to perform DNA testing.

DNA testing can also be performed with newborns. The blood samples used for screening for PKU and other disorders contain enough of the baby's blood cells to permit DNA analysis using PCR-based techniques. Screening tests using DNA analysis are now being used for sickle-cell disease and cystic fibrosis; similar tests for other diseases will surely follow.

Of the numerous methods of DNA testing available, two are the most widespread. We will describe their use to detect the mutation in the β -globin gene that results in sickle-cell disease.

SCREENING FOR ALLELE-SPECIFIC CLEAVAGE DIFFERENCES The first method uses RFLP analysis, as we described earlier. There is a difference between the normal and the sickle allele of the β -globin gene, with respect to a restriction enzyme recognition sequence. Around the sixth codon in the normal gene is the sequence



This sequence is recognized by the restriction enzyme *Mst*II, which will cleave DNA at



where N is any base. In the sickle allele, the DNA sequence is



The point mutation at codon 6 makes this sequence unrecognizable by *Mst*II. The sequence surrounding the mutant site can be amplified by PCR and digested with *Mst*II. Gel electrophoresis is used to distinguish between PCR products derived from the normal allele, which are cut by the enzyme, and products from the sickle allele, which are not cut (**Figure 15.20**).

This *allele-specific cleavage* method of DNA testing works only if a restriction enzyme exists that can recognize the sequence of either the normal or the mutant allele.

SCREENING BY ALLELE-SPECIFIC OLIGONUCLEOTIDE HYBRIDIZATION

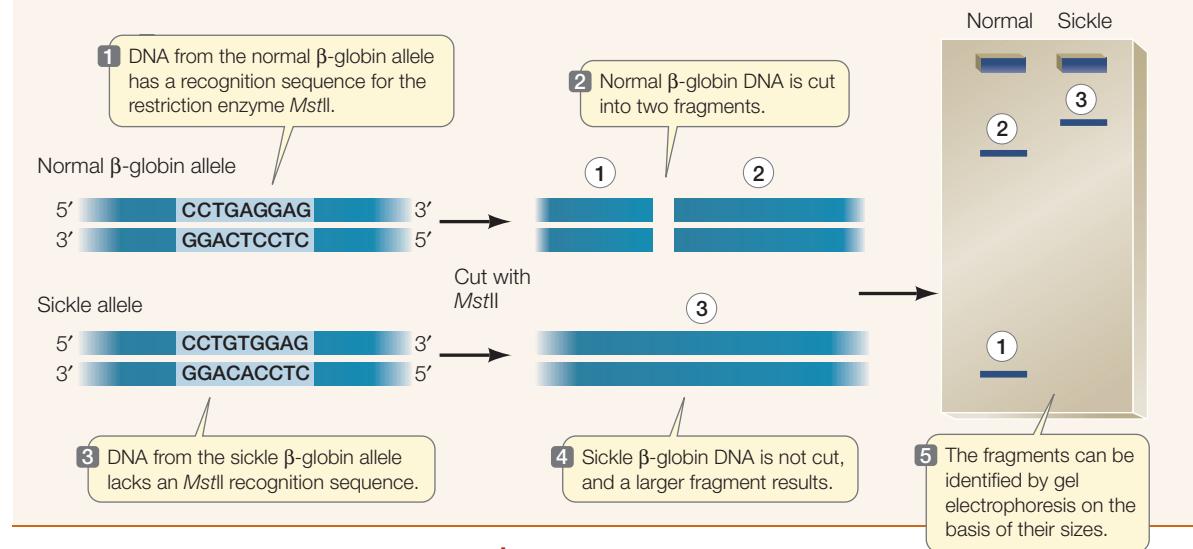
The *allele-specific oligonucleotide hybridization* method uses short synthetic DNA strands called *oligonucleotide probes* that will hybridize with denatured PCR products from either the normal or the mutant allele. Usually, an oligonucleotide probe of at least a dozen bases is needed to form a stable double helix with the target DNA. If the probe is radioactively or fluorescently labeled, hybridization can be readily detected (**Figure 15.21**).

Detection of a mutation by either DNA screening method can be used for diagnosis of a genetic disease, so that appropriate treatment can begin. In addition, identification provides a person with important information about his or her genome.

TOOLS FOR INVESTIGATING LIFE

15.20 DNA Testing by Allele-Specific Cleavage

Allele-specific cleavage can be used to detect mutations such as the one that causes sickle-cell disease.



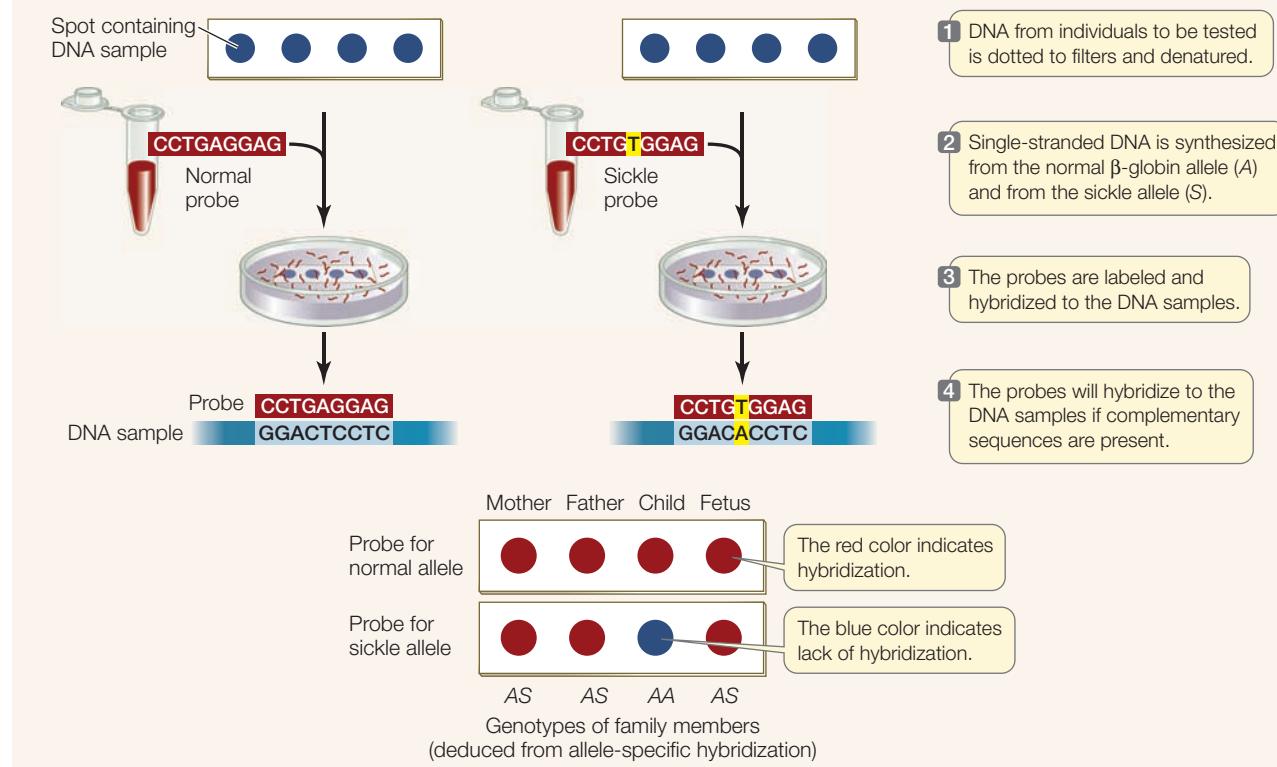
yourBioPortal.com

GO TO Animated Tutorial 15.2 • DNA Testing by Allele-Specific Cleavage

TOOLS FOR INVESTIGATING LIFE

15.21 DNA Testing by Allele-Specific Oligonucleotide Hybridization

Testing of this family reveals that three of them are heterozygous carriers of the sickle allele. The first child, however, has inherited two normal alleles and is neither affected by the disease nor a carrier.



15.5 RECAP

Genetic screening can be used to identify people who have, are predisposed to, or are carriers of, genetic diseases. Screening can be done at the phenotype level by identifying an abnormal protein such as an enzyme with altered activity. It can also be done at the genotype level by direct testing of DNA.

- How are newborn babies screened for PKU?
See p. 334 and Figure 15.19
- What is the advantage of screening for genetic mutations by allele-specific oligonucleotide hybridization relative to screening for allele-specific cleavage differences? See p. 335 and Figures 15.20 and 15.21

Ongoing research has resulted in the development of increasingly accurate diagnostic tests and a better understanding of various genetic diseases at the molecular level. This knowledge is now being applied to the development of new treatments for genetic diseases. In the next section we will survey various approaches to treatment, including modifications of the mutant phenotype and gene therapy, in which the normal version of a mutant gene is supplied.

15.6 How Are Genetic Diseases Treated?

Most treatments for genetic diseases simply try to alleviate the patient's symptoms. But to effectively treat these diseases—whether they affect all cells, as in inherited disorders such as PKU, or only somatic cells, as in cancer—physicians must be able to diagnose the disease accurately, understand how the disease works at the molecular level, and intervene early, before the disease ravages or kills the individual. There are two main approaches to treating genetic diseases: modifying the disease phenotype, or replacing the defective gene.

Genetic diseases can be treated by modifying the phenotype

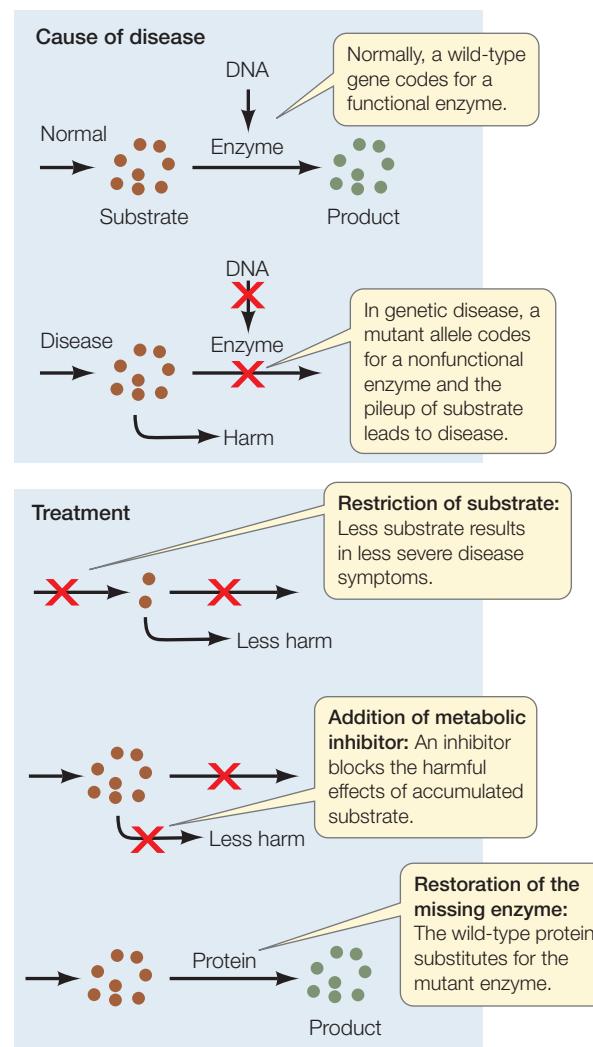
Altering the phenotype of a genetic disease so that it no longer harms an individual is commonly done in one of three ways: by restricting the substrate of a deficient enzyme, by inhibiting a harmful metabolic reaction, or by supplying a missing protein product (Figure 15.22).

RESTRICTING THE SUBSTRATE Restricting the substrate of a deficient enzyme is the approach taken when a newborn is diagnosed with PKU. In this case, the deficient enzyme is phenylalanine hydroxylase, and the substrate is phenylalanine. The infant's inability to break down phenylalanine in food leads to a buildup of the substrate, which causes the clinical symptoms. So the infant is immediately put on a special diet that contains only enough phenylalanine for immediate use. Lofenelac, a milk-based product that is low in phenylalanine, is fed to these

infants just like formula. Later, certain fruits, vegetables, cereals, and noodles low in phenylalanine can be added to the diet. Meat, fish, eggs, dairy products, and bread, which contain high amounts of phenylalanine, must be avoided, especially during childhood, when brain development is most rapid. The artificial sweetener aspartame must also be avoided because it is made of two amino acids, one of which is phenylalanine.

People with PKU are generally advised to stay on a low-phenylalanine diet for life. Although maintaining these dietary restrictions may be difficult, it is effective. Numerous follow-up studies since newborn screening was initiated have shown that people with PKU who stay on the diet are no different from the rest of the population in terms of mental ability. This is an impressive achievement in public health, given the severity of mental retardation in untreated patients.

METABOLIC INHIBITORS In Section 11.7, we described how drugs that are inhibitors of various cell cycle processes are used to treat cancer. Drugs are also used to treat the symptoms of many genetic diseases. As biologists have gained insight into the molec-



15.22 Strategies for Treating Genetic Diseases

ular characteristics of these diseases and the specific proteins involved, a more specific approach to treatment is taking shape. This is called *molecular medicine*.

An example of this approach is the treatment of chronic myelogenous leukemia. In this cancer, certain white blood cells undergo a gain-of-function mutation, making a totally new protein that is not made in any other cells. This new protein was isolated and a drug was made that specifically targets and inactivates the protein, thereby preventing the proliferation of the cancerous cells. The result has been greatly improved survival in these patients.

SUPPLYING THE MISSING PROTEIN An obvious way to treat a disease phenotype in which a functional protein is missing is to supply that protein. This approach is the basis of treatment for hemophilia, in which the missing blood factor VIII is supplied to the patient. At first this protein was obtained from blood and was sometimes contaminated with viruses or other pathogens. Now, however, the production of human clotting proteins by recombinant DNA technology (see Chapter 18) has made it possible to provide the protein in a much purer form.

Unfortunately, the phenotypes of many diseases caused by genetic mutations are very complex. In these cases, simple interventions like those we have just described do not work. Indeed, a recent survey of 351 diseases caused by single-gene mutations showed that current therapies increased patients' life spans by only 15 percent.

Gene therapy offers the hope of specific treatments

Clearly, if a cell lacks a functional allele, it would be optimal to provide that allele. This is the aim of gene therapy. Diseases ranging from rare inherited disorders caused by single-gene mutations to cancer are under intensive investigation, in an effort to develop gene therapy treatments.

The object of **gene therapy** is to insert a new gene that will be expressed in the host. The new DNA must be attached to a promoter that will be active in human cells. The physicians who are developing such treatments are confronted by numerous challenges. They must find an effective way for the new gene to be taken up by the patient's cells, for the gene to be precisely inserted into the host DNA, and for the gene to be expressed.

Which human cells should be the targets of gene therapy? The best approach would be to replace the nonfunctional allele with a functional one in every cell of the body. But delivery of a gene to every cell poses a formidable challenge. Until recently, attempts at gene therapy have used *ex vivo* techniques. That is, physicians have taken cells from the patient's body, added the new gene to those cells in the laboratory, and then returned the cells to the patient in the hope that the correct gene product would be made (Figure 15.23). A successful example demonstrates this technique.

Adenosine deaminase is needed for the maturation of white blood cells, and a genetic disease results when a person is homozygous for a mutant allele for this enzyme. People without this enzyme have severe immune system deficiencies. The wild-type gene for adenosine deaminase has been isolated and in-

INVESTIGATING LIFE

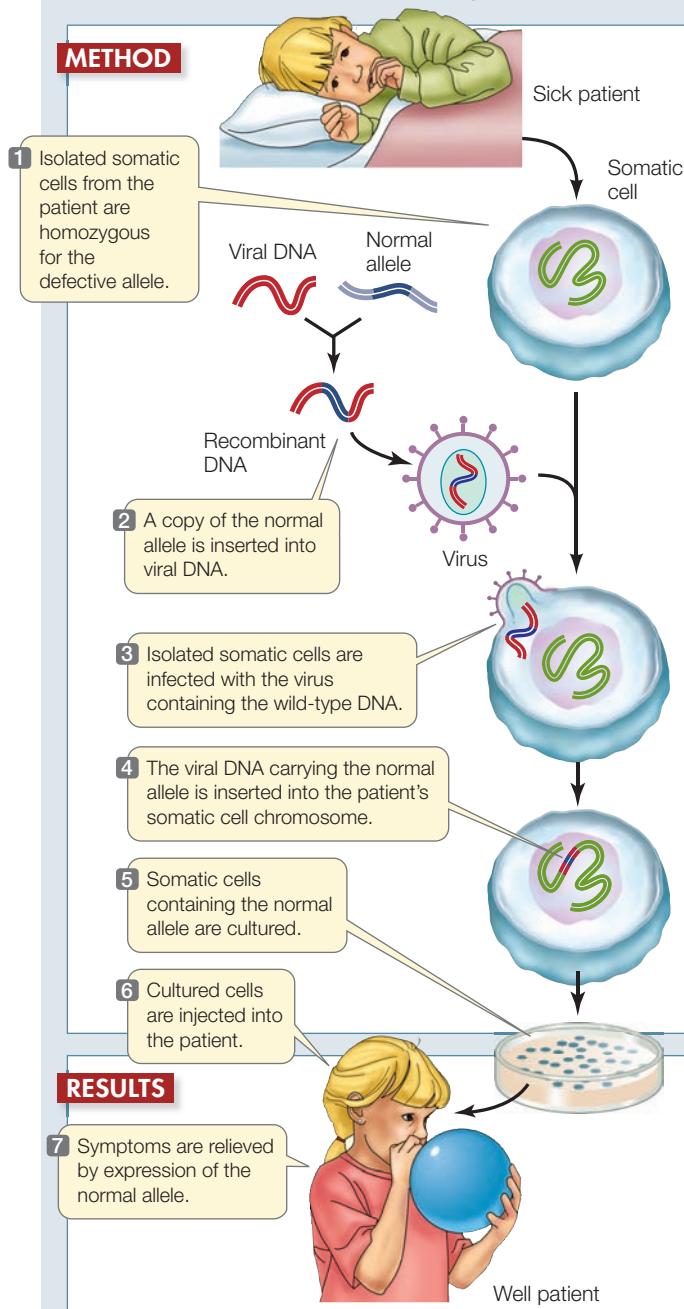
15.23 Gene Therapy: The Ex Vivo Approach

New genes are added to somatic cells taken from a patient's body. These cells are then returned to the body to make the missing gene product.

HYPOTHESIS

The introduction and expression of a normal allele can help a patient with two defective alleles for an important gene.

METHOD



RESULTS

CONCLUSION

Gene therapy can be effective in relieving symptoms caused by a genetic disease.

Go to yourBioPortal.com for original citations, discussions, and relevant links for all INVESTIGATING LIFE figures.

serted into a virus that can carry the gene into white blood cells of a patient lacking the enzyme. The recombinant virus lacks the genes for reproduction inside cells, but retains the genes coding for cell uptake and insertion into the host DNA. The recombinant virus was added to white blood cells from a patient that had inherited the mutant form of adenosine deaminase. The wild-type adenosine deaminase gene became inserted into the cells' chromosomes, along with viral DNA. When these transformed white blood cells with the wild-type gene were put back into the patient, the cells made adenosine deaminase and the patient's condition improved.

The other approach to gene therapy is to insert the gene directly into cells in the body of the patient. This *in vivo* approach is being attempted for various types of cancer. Lung cancer cells, for example, are accessible to such treatment if the DNA is given as an aerosol through the respiratory system. Several thousand patients, over half of them with cancer, have undergone this treatment. In preliminary clinical trials, people are given the therapy to see whether it has any toxicity and whether the new gene is actually incorporated into the patients' genomes. In more ambitious trials, larger numbers of patients receive the therapy with the hope that their disease will disappear, or at least improve.

15.6 RECAP

Treatment of a human genetic disease may involve an attempt to modify the abnormal phenotype by restricting the substrate of a deficient enzyme, inhibiting a harmful metabolic reaction, or supplying a missing protein. On the other hand, gene therapy aims to address a genetic defect by inserting a normal allele into a patient's cells.

- How do metabolic inhibitors used in chemotherapy function in treating cancer? [See pp. 337–338 and Figure 5.22](#)
- How does ex vivo gene therapy work? Can you give an example? [See p. 338 and Figure 15.23](#)

In this chapter, we dealt with mutations in general, focusing on DNA changes that affect phenotypes through specific protein products. But there is much more to molecular genetics than genes and proteins. Determining which genes will be expressed when and where is a major function of the genome. In Chapter 16 we turn to gene regulation.

CHAPTER SUMMARY

15.1 What Are Mutations?

- **Mutations** are heritable changes in DNA. **Somatic mutations** are passed on to daughter cells, but only **germ line mutations** are passed on to sexually produced offspring.
- **Point mutations** result from alterations in single base pairs of DNA. **Silent mutations** can occur in noncoding DNA or in coding regions of genes and do not affect the amino acid sequences of proteins. **Missense, nonsense, and frame-shift** mutations all cause changes in protein sequences. [Review Figure 15.2](#)
- Chromosomal mutations (**deletions, duplications, inversions, or translocations**) involve large regions of chromosomes. [Review Figure 15.4](#)
- **Spontaneous mutations** occur because of instabilities in DNA or chromosomes. **Induced mutations** occur when a mutagen damages DNA. [Review Figure 15.5](#)
- Mutations can occur in hot spots where cytosine has been methylated to 5-methylcytosine. [Review Figure 15.6](#)
- Mutations, although often detrimental to an individual organism, are the raw material of evolution.

15.2 How Are DNA Molecules and Mutations Analyzed?

- **Restriction enzymes**, which are made by microorganisms as a defense against viruses, bind to and cut DNA at specific **recognition sequences** (also called **restriction sites**). These enzymes can be used to produce small fragments of DNA for study, a technique known as restriction digestion. [Review Figure 15.7](#)
- DNA fragments can be separated by size using **gel electrophoresis**. [Review Figure 15.8, ANIMATED TUTORIAL 15.1](#)
- **DNA fingerprinting** is used to distinguish between specific individuals, or to reveal which individuals are most closely related

to one another. It involves the detection of DNA polymorphisms, including **single nucleotide polymorphisms (SNPs)** and **short tandem repeats (STRs)**. [Review Figure 15.9](#)

- The goal of the DNA barcoding project is to sequence a single region of DNA in all species for identification purposes.

15.3 How Do Defective Proteins Lead to Diseases?

- Abnormalities in nearly all classes of proteins, including enzymes, transport proteins, receptor proteins, and structural proteins, have been implicated in genetic diseases.
- While a single amino acid difference can be the cause of disease, amino acid variations have been detected in many functional proteins. [Review Figure 15.13](#)
- Transmissible spongiform encephalopathies (TSEs) are degenerative brain diseases that can be transmitted from one animal to another by consumption of infected tissues. The infective agent is a **prion**, a protein with an abnormal conformation.
- **Multifactorial** diseases are caused by the interactions of many genes and proteins with the environment. They are much more common than diseases caused by mutations in a single gene.
- Predictable patterns of inheritance are associated with some human genetic diseases. Autosomal recessive, autosomal dominant, and sex-linked patterns are common.

15.4 What DNA Changes Lead to Genetic Diseases?

- It is possible to isolate both the mutant genes and the abnormal proteins responsible for human diseases. [Review Figure 15.15, WEB ACTIVITY 15.1](#)
- The effects of fragile-X syndrome worsen with each generation. This pattern is the result of an **expanding triplet repeat**. [Review Figure 15.18](#)

15.5 How Is Genetic Screening Used to Detect Human Diseases?

- **Genetic screening** is used to detect human genetic diseases, alleles predisposing people to those diseases, or carriers of those diseases.
- Genetic screening can be done by looking for abnormal protein expression. **Review Figure 15.19**
- **DNA testing** is the direct identification of mutant alleles. Any cell can be tested at any time in the life cycle.
- The two predominant methods of DNA testing are the allele-specific cleavage method and allele-specific oligonucleotide

hybridization method. **Review Figures 15.20 and 15.21, ANIMATED TUTORIAL 15.2**

15.6 How Are Genetic Diseases Treated?

- There are three ways to modify the phenotype of a genetic disease: restrict the substrate of a deficient enzyme, inhibit a harmful metabolic reaction, or supply a missing protein. **Review Figure 15.22**
- Cancer is treated with metabolic inhibitors.
- In **gene therapy**, a mutant gene is replaced with a normal gene. Both ex vivo and in vivo therapies are being developed. **Review Figure 15.23**

SELF-QUIZ

1. Phenylketonuria is an example of a genetic disease in which
 - a single enzyme is not functional.
 - inheritance is sex-linked.
 - two parents without the disease cannot have a child with the disease.
 - mental retardation always occurs, regardless of treatment.
 - a transport protein does not work properly.
2. Mutations of the gene for β -globin
 - are usually lethal.
 - occur only at amino acid position 6.
 - number in the hundreds.
 - always result in sickling of red blood cells.
 - can always be detected by gel electrophoresis.
3. Multifactorial (complex) diseases
 - are less common than single-gene diseases.
 - involve the interaction of many genes with the environment.
 - affect less than 1 percent of humans.
 - involve the interactions of several mRNAs.
 - are exemplified by sickle-cell disease.
4. In fragile-X syndrome,
 - females are affected more severely than males.
 - a short sequence of DNA is repeated many times to create the fragile site.
 - both the X and Y chromosomes tend to break when prepared for microscopy.
 - all people who carry the gene that causes the syndrome are mentally retarded.
 - the basic pattern of inheritance is autosomal dominant.
5. Most genetic diseases are rare because
 - each person is unlikely to be a carrier for harmful alleles.
 - genetic diseases are usually sex-linked and so uncommon in females.
 - genetic diseases are always dominant.
 - two parents probably do not carry the same recessive alleles.
 - mutation rates in humans are low.
6. Mutational “hot spots” in human DNA
 - always occur in genes that are transcribed.
 - are common at cytosines that have been modified to 5-methylcytosine.
 - involve long stretches of nucleotides.
 - occur only where there are long repeats.
 - are very rare in genes that code for proteins.
7. Newborn genetic screening for PKU
 - is very expensive.
 - detects phenylketones in urine.
 - has not led to the prevention of mental retardation resulting from this disorder.
 - should be done during the second or third day of an infant’s life.
 - uses bacterial growth to detect excess phenylketones in blood.
8. Genetic diagnosis by DNA testing
 - detects only mutant and not normal alleles.
 - can be done only on eggs or sperm.
 - involves hybridization to rRNA.
 - often utilizes restriction enzymes and a polymorphic site.
 - cannot be done with PCR.
9. Which of the following is *not* a way to treat a genetic disease?
 - Inhibiting a harmful biochemical reaction
 - Adding the wild-type allele to cells expressing the mutation
 - Restricting the substrate of a harmful biochemical reaction
 - Replacing a mutant allele with the wild-type allele in the fertilized egg
 - Supplying a wild-type protein that is missing due to mutation
10. Current treatments for genetic diseases include all of the following *except*
 - restricting a dietary substrate.
 - replacing the mutant gene in all cells.
 - alleviating the patient’s symptoms.
 - inhibiting a harmful metabolic reaction.
 - supplying a protein that is missing.

FOR DISCUSSION

1. In the past, it was common for people with phenylketonuria (PKU) who were placed on a low-phenylalanine diet after birth to be allowed to return to a normal diet during their teenage years. Although the levels of phenylalanine in their blood were high, their brains were thought to be beyond the stage when they could be harmed. If a woman with PKU becomes pregnant, however, a problem arises. Typically, the fetus is heterozygous, but is unable, at early stages of development, to metabolize the high levels of phenylalanine that arrive from the mother's blood. Why is the fetus likely to be heterozygous? What do you think would happen to the fetus during this "maternal PKU" situation? What would be your advice to a woman with PKU who wants to have a child?
2. Cystic fibrosis is an autosomal recessive disease in which thick mucus is produced in the lungs and airways. The gene responsible for this disease encodes a protein composed of 1,480 amino acids. In most patients with cystic fibrosis, the protein has 1,479 amino acids: a phenylalanine is missing at position 508. A baby is born with cystic fibrosis. He has an older brother who is not affected. How would you test the DNA of the older brother to determine whether he is a carrier for cystic fibrosis? How would you design a gene therapy protocol to "cure" the cells in the younger brother's lungs and airways?
3. A number of efforts are under way to identify human genetic polymorphisms that correlate with multifactorial diseases such as diabetes, heart disease, and cancer. What would be the uses of such information? What concerns do you think are being raised about this kind of genetic testing?

ADDITIONAL INVESTIGATION

Tay-Sachs disease is caused by a recessively inherited mutation in the gene coding for the enzyme hexosaminidase A (HexA), which normally breaks down a lipid called GM2 ganglioside. Accumulation of this lipid in the brain leads to progressive deterioration of the nervous system and death, usually by age 4. HexA activity in blood serum is 0–6 percent in homozygous recessives and 7–35 percent in heterozygous carriers, compared

to non-carriers (100 percent). The most common mutation in the *HexA* gene is an insertion of four base pairs, which presumably leads to a premature stop codon. How would you do genetic screening for carriers of this disease by enzyme testing and by DNA testing? What are the advantages of DNA testing? How would you investigate the premature stop codon hypothesis?

WORKING WITH DATA (GO TO yourBioPortal.com)

Gene Therapy: The Ex Vivo Approach In this exercise, you use the original research paper to examine the protocol used to treat two patients with gene therapy for adenosine deaminase deficiency (Figure 15.23). You will examine the kinds of evi-

dence used to detect the wild-type gene in the cells of these patients, and will analyze the results in terms of immune system cell function.

16

Regulation of Gene Expression

Alcoholism and the control of gene expression

Many people drink alcoholic beverages but relatively few of them become addicted (alcoholic). When they do, the results are often disastrous, both socially and physiologically. Alcoholism often disrupts relationships with family, friends, and colleagues. Lost productivity leads to economic costs estimated at over \$100 billion per year in the U.S. alone. Physiologically, alcoholism is characterized by a compulsion to consume alcohol, tolerance (increasing doses are needed for the same effect), and dependence (abrupt cessation of consumption leads to severe withdrawal symptoms). In most of these people, alcohol acts not just to provide pleasant sensations (positive reinforcement) but also to alleviate unpleasant ones such as anxiety (negative reinforcement).

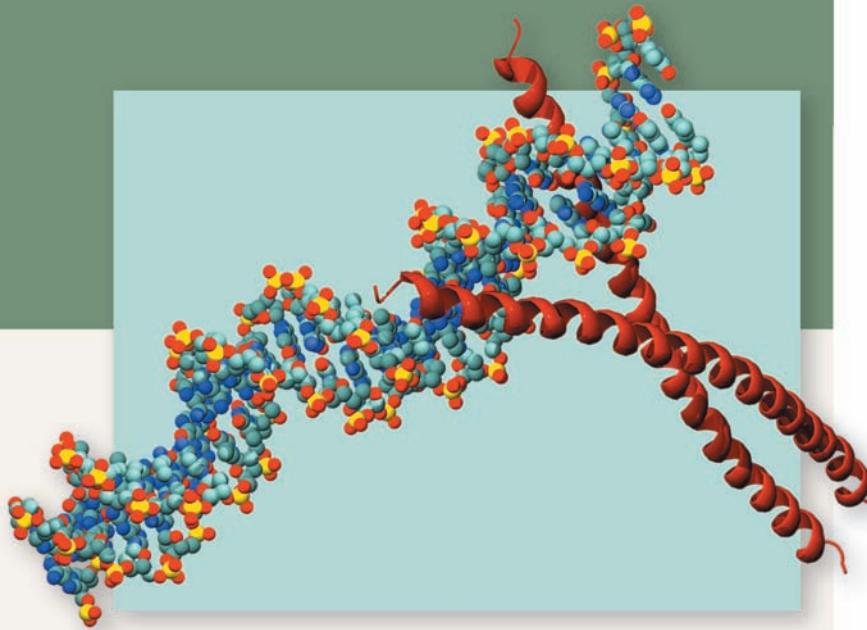


Why do only some people become alcoholic? Alcoholism is a complex behavioral disease. Psychologists sometimes speak of "addictive personalities," and genetic studies indicate there may be inherited factors. It would help both alcoholics and those who treat them if we understood the differences in brain chemistry between alcoholic and nonalcoholic individuals. But we can't do the necessary experiments on humans; instead, animal models are used to study alcoholism at the molecular level. James Murphy at Indiana University has bred a strain of rats, called P rats, that prefer alcohol when given the choice of alcohol-containing or alcohol-free water. These rats show many of the symptoms of true addiction, including compulsive drinking, tolerance, and withdrawal. In effect, they are a genetic strain of alcoholic animals.

People often drink alcoholic beverages to relieve anxiety, and there are clear links between anxiety disorders and alcoholism. Like many of their human counterparts, the P rats appear more anxious than wild-type rats, spending more time in a closed rather than an open environment. Drinking alcohol alters this behavior and seems to relieve their anxiety.

There may also be a link between the transcription factor CREB and alcohol consumption. CREB (or cyclic AMP response element binding protein) is especially abundant in the brain and regulates the expression of hundreds of genes that are important in metabolism. CREB becomes activated when it is phosphorylated by the enzyme protein kinase A, which in turn is activated by the second messenger cyclic AMP. In an effort to understand the molecular basis of alcoholism and anxiety, neuroscientist Subhash Pandey and his colleagues at the University of Illinois compared CREB levels in the brains of P rats and wild-type rats.

Alcoholism Huge social and economic costs are associated with alcohol abuse. Scientists are trying to understand its molecular basis.



An Explanation for Alcoholism? The transcription factor, CREB, binds to DNA and activates promoters of genes involved in addictive behaviors in alcoholism.

They found that P rats have inherently lower levels of CREB in certain parts of the brain. When these rats consumed alcohol, the total levels of CREB did not increase, but the levels of phosphorylated CREB did. It is the phosphorylated version of CREB that binds to DNA and regulates gene transcription.

The prospect that CREB, a transcription factor that regulates gene expression, is a key element in the genetic propensity for alcoholism is important because it begins to explain the molecular nature of a complex behavioral disease. Such understanding may permit more effective treatment of alcohol abuse or its prevention. Equally important to our purpose here, it underscores the importance of the regulation of gene expression in biological processes.

IN THIS CHAPTER we will focus on the control of gene expression in many types of organisms. We begin with the simplest systems, viruses, which undertake an ordered series of molecular events when they infect a host cell. Then we turn to prokaryotes, which respond to changes in their environment with coordinated changes in gene expression. In eukaryotes, similar principles are used to regulate gene expression, but with added levels of complexity. Finally, we turn to the regulation of gene expression by modification of the genome—the field of epigenetics.

CHAPTER OUTLINE

- 16.1 How Do Viruses Regulate Their Gene Expression?
- 16.2 How Is Gene Expression Regulated in Prokaryotes?
- 16.3 How Is Eukaryotic Gene Transcription Regulated?
- 16.4 How Do Epigenetic Changes Regulate Gene Expression?
- 16.5 How Is Eukaryotic Gene Expression Regulated After Transcription?

16.1 How Do Viruses Regulate Their Gene Expression?

“A virus is a piece of bad news wrapped in protein.” This quote from immunologist Sir Peter Medawar is certainly true for the cells that viruses infect. As we describe in Chapter 13, a virus injects its genetic material into a host cell and turns that cell into a virus factory (see Figure 13.3). Viral life cycles are very efficient. Perhaps the record is held by poliovirus: a single poliovirus infecting a mammalian cell can produce over 100,000 new virus particles!

Unlike organisms, **viruses** are *acellular*; that is, they are not cells, do not consist of cells, and do not carry out many of the processes characteristic of life. Most virus particles, called **virions**, are composed of only nucleic acid and a few proteins. Viruses do not carry out two of the basic functions of cellular life: they do not regulate the transport of substances into and out of themselves by membranes, and they do not perform metabolic functions involved with taking in nutrients, refashioning them, and expelling wastes. But they can reproduce in systems that do perform these metabolic functions—namely, living cells. By studying the relatively simple viral reproductive cycle, biologists have discovered principles of gene expression and its regulation that apply to cellular systems that may be much more complex.

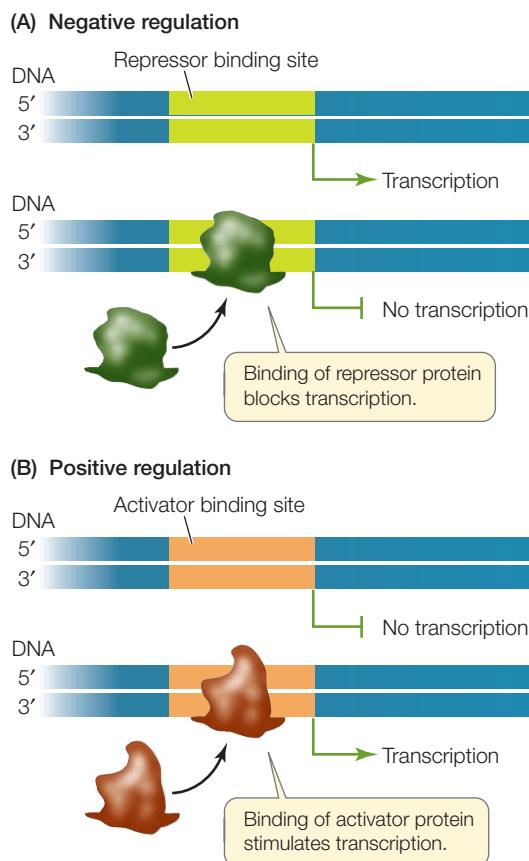
As we describe in Chapter 14, gene expression begins at the *promoter*, where RNA polymerase binds to initiate transcription. In a genome with many genes, not all promoters are active at a given time—there is *selective gene transcription*. The “decision” regarding which genes to activate involves two types of regulatory proteins that bind to DNA: repressor proteins and activator proteins. In both cases, these proteins bind to the promoter to regulate the gene (**Figure 16.1**):

- In **negative regulation**, the gene is normally transcribed. Binding of a repressor protein prevents transcription.
- In **positive regulation**, the gene is normally not transcribed. An activator protein binds to stimulate transcription.

You will see these mechanisms, or combinations of them, as we examine regulation in viruses, prokaryotes, and eukaryotes.

Bacteriophage undergo a lytic cycle

The Hershey–Chase experiment (see Figure 13.4) involved the typical viral reproductive cycle, the **lytic cycle**, so named because the infected host cell lyses (bursts), releasing progeny viruses. Once a virus has injected its nucleic acid into a cell, that

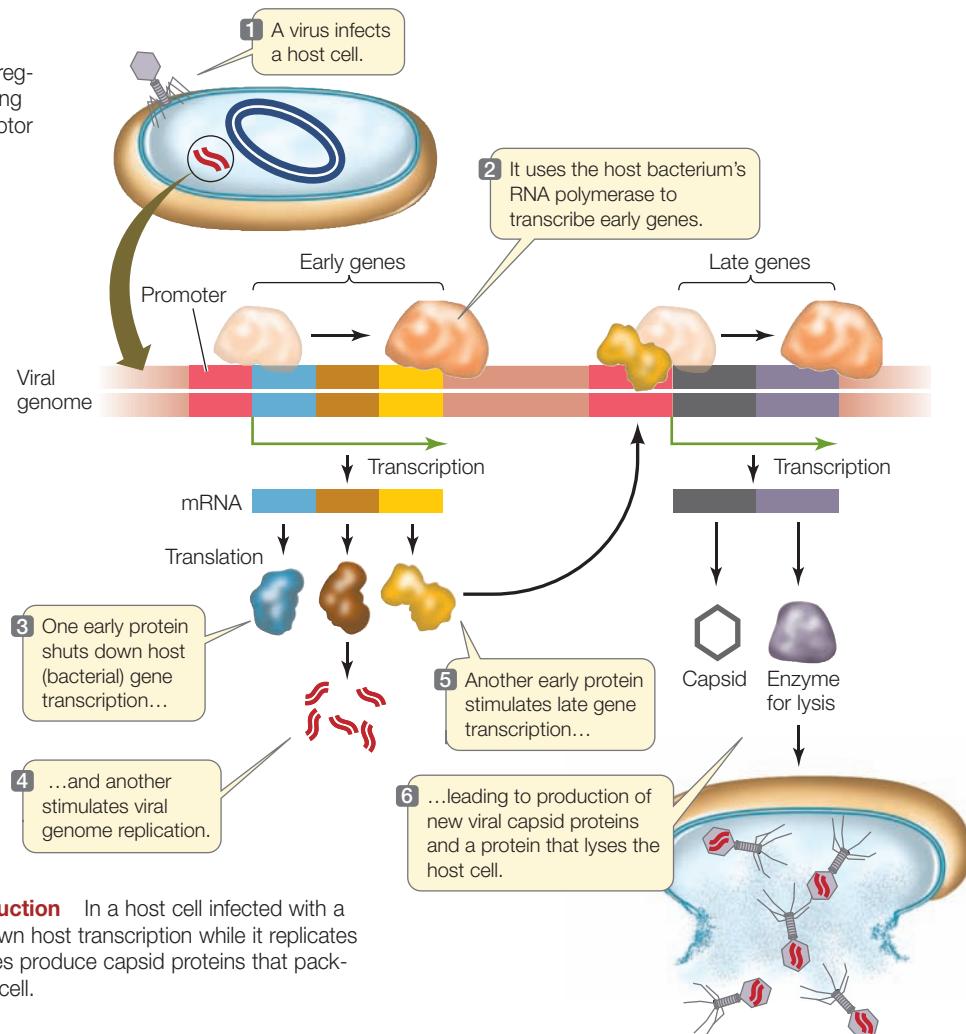
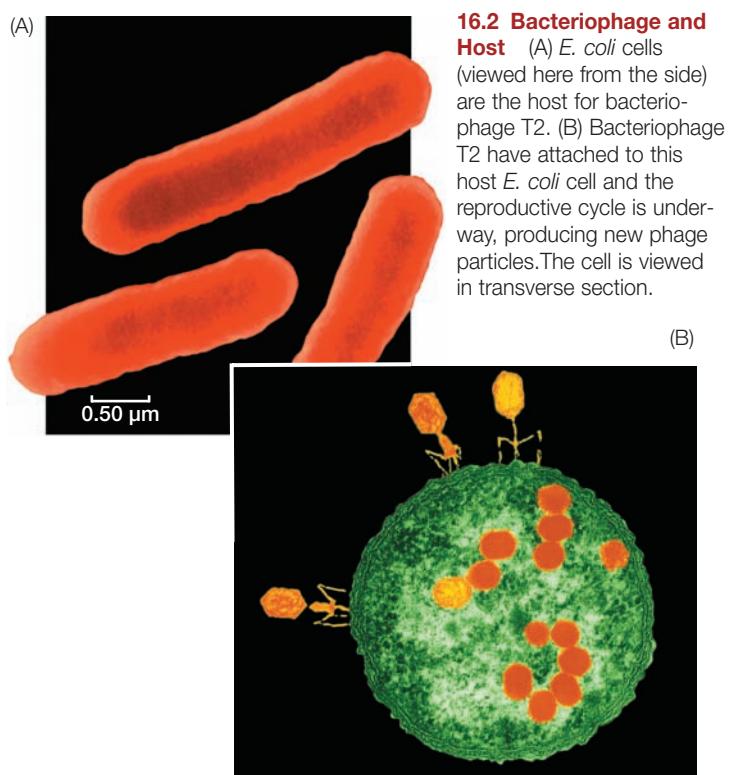


16.1 Positive and Negative Regulation Proteins regulate gene expression by binding to DNA and preventing or allowing RNA polymerase to bind DNA at the promotor region to control transcription.

nucleic acid takes over the host's synthetic machinery. In the case of some **bacteriophage** (viruses that infect bacteria), the process is extremely rapid—within 15 minutes, new phage particles appear in the bacterial cell (**Figure 16.2**). Ten minutes later, the “game is over,” and these particles are released from the lysed cell. What happens during this rapid life cycle?

At the molecular level, the reproductive cycle of a lytic virus has two stages: early and late (**Figure 16.3**). Look for both positive and negative regulation of gene expression in these events:

- The viral genome contains a promoter that binds host RNA polymerase. In the *early stage* (1–2 minutes after phage DNA entry) viral genes that lie adjacent



16.3 The Lytic Cycle: A Strategy for Viral Reproduction In a host cell infected with a virus, the viral genome uses its early genes to shut down host transcription while it replicates itself. Once the viral genome is replicated, its late genes produce capsid proteins that package the genome and other proteins that lyse the host cell.

to this promoter are transcribed. These early genes often encode proteins that shut down host transcription and stimulate viral genome replication and transcription of viral late genes. Three minutes after DNA entry, viral nucleic acid enzymes digest the host's chromosome, providing nucleotides for the synthesis of viral genomes.

- In the *late stage*, viral late genes are transcribed; they encode the viral capsid proteins and enzymes that lyse the host cell to release the new virions. This begins 9 minutes after DNA entry and 6 minutes before the first new phage particles appear.

The whole process—from binding and infection to release of new phage—takes about half an hour. During this period, the sequence of transcriptional events is carefully controlled to produce complete, infective viroids.

Some bacteriophage can carry bacterial genes from one cell to another

During the lytic cycle some bacteriophage package their DNA in **capsids** (outer shells). In rare cases, a bacterial DNA fragment is inserted into a capsid instead of, or along with, the phage DNA. When such a virion infects another bacterium, the bacterial DNA is injected into the new host cell, a mechanism of gene transfer called **transduction**. The viral infection does not produce new viruses. Instead, the incoming DNA fragment can re-

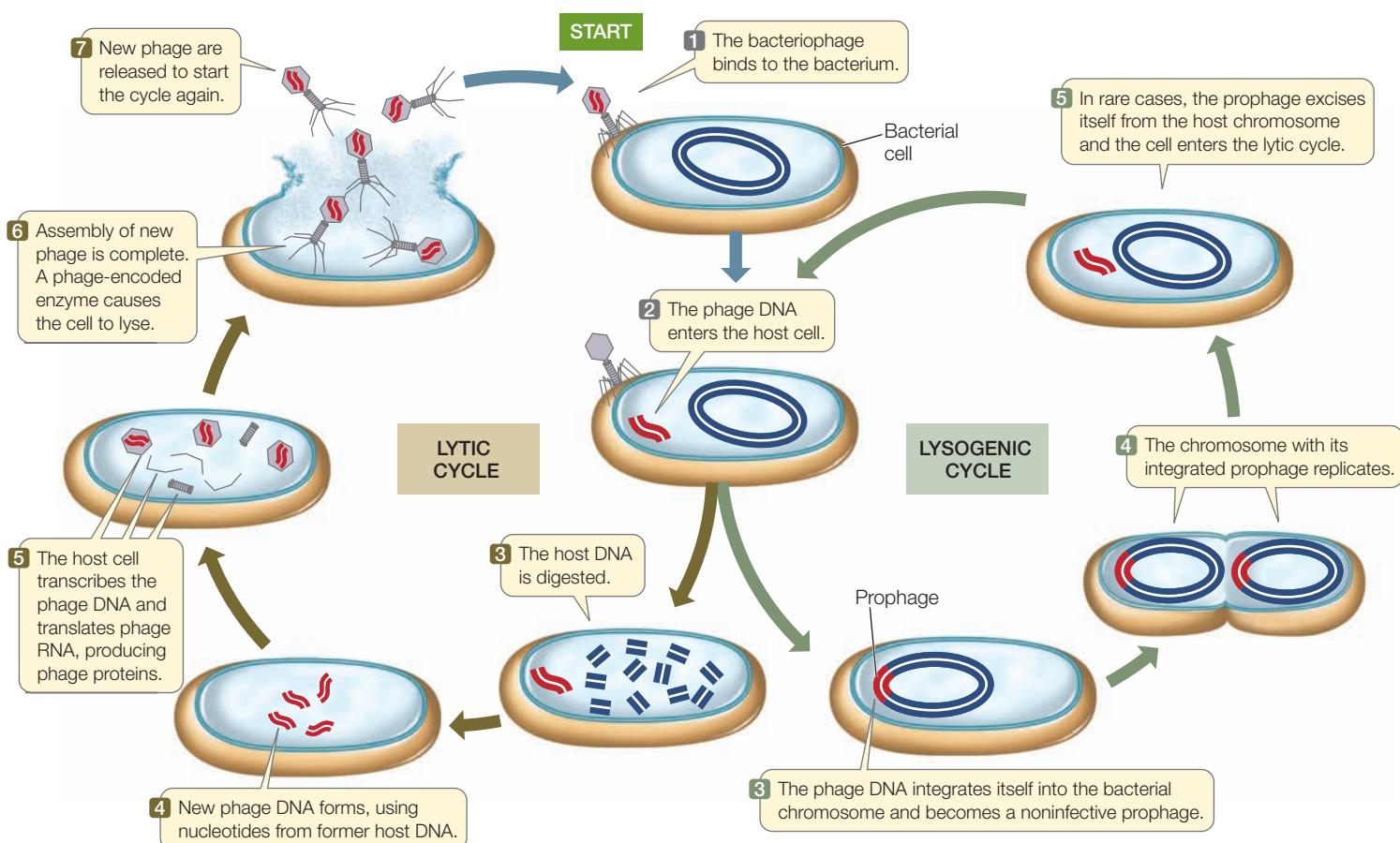
combine with the host chromosome, replacing host genes with genes from the virus's former host. The recipient cell survives under these conditions because there is no virus replication.

Some bacteriophage can undergo a lysogenic cycle

Like all nucleic acid genomes, those of viruses can mutate and evolve by natural selection. Some viruses have evolved an advantageous process called **lysogeny** that postpones the lytic cycle. In lysogeny, the viral DNA becomes integrated into the host DNA and becomes a **prophage** (Figure 16.4). As the host cell divides, the viral DNA gets replicated along with that of the host. The prophage can remain inactive within the bacterial genome for thousands of generations, producing many copies of the original viral DNA.

However, if the host cell is not growing well, the virus “cuts its losses.” It immediately switches to a lytic cycle, in which the prophage excises itself from the host chromosome and reproduces. In other words, the virus is able to enhance its chances of multiplication and survival by inserting its DNA into the host chromosome, where it sits as a silent partner until conditions are right for lysis.

16.4 The Lytic and Lysogenic Cycles of Bacteriophage In the lytic cycle, infection of a bacterium by viral DNA leads directly to the multiplication of the virus and lysis of the host cell. In the lysogenic cycle, an inactive prophage is integrated into the host DNA where it is replicated during the bacterial life cycle.



16.5 Control of Bacteriophage λ Lysis and Lysogeny

Two regulatory proteins, Cro and cl, compete to control expression of one another and genes for viral lysis and lysogeny.

Uncovering the regulation of gene expression that underlies the lysis/lysogeny switch was a major achievement of molecular biologists. Here we present just an outline of the process to give you an idea of the positive and negative regulatory mechanisms involved (Figure 16.5). The model virus *bacteriophage λ* (lambda) has been used extensively to study the lysogenic mechanism.

How does the phage “know” when to switch to the lytic cycle? A kind of “genetic switch” senses conditions within the host. Two viral regulatory proteins, cl and Cro, compete for two promoters on the phage DNA (Figure 16.5). These two promoters control the transcription of the viral genes involved in the lytic and the lysogenic cycles, respectively, and the two regulatory proteins have opposite effects on each promoter.

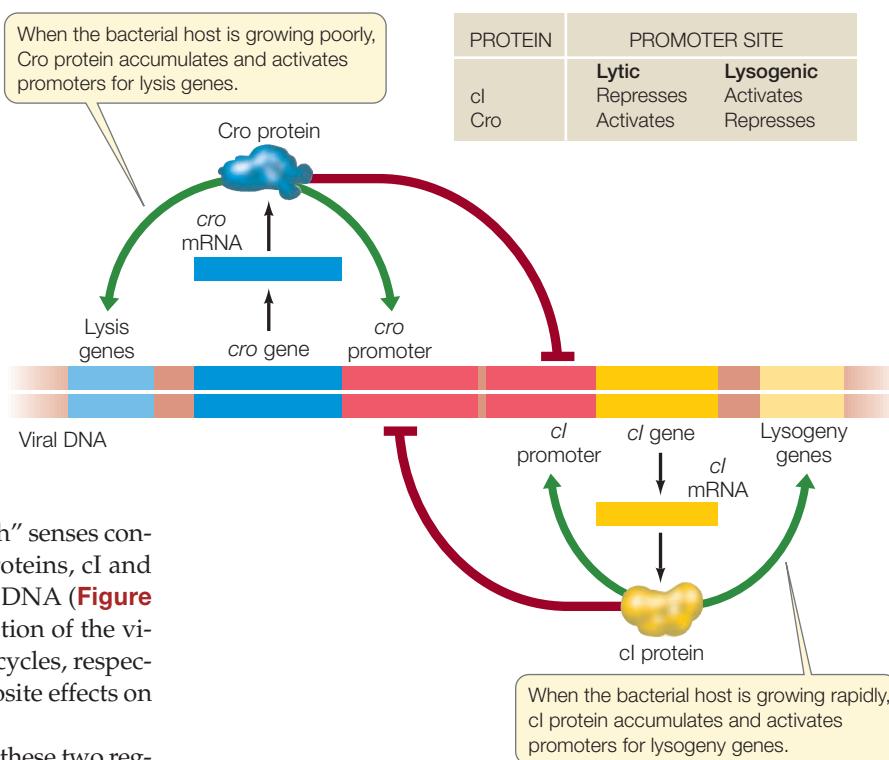
Phage infection is essentially a “race” between these two regulatory proteins. In a rapidly growing *E. coli* host cell, Cro synthesis is low, so cl “wins,” and the phage enters a lysogenic cycle. If the host cell is growing slowly, Cro synthesis is higher, and the genes involved in lysis are activated. The two regulatory proteins are made very early in phage infection, and each binds to a specific DNA sequence.

The reproductive cycle of bacteriophage λ is a paradigm for our understanding of viral life cycles in general. This relatively simple system has served as a model to help us understand how the complicated reproductive cycles of other viruses, including HIV, are controlled.

Eukaryotic viruses have complex regulatory mechanisms

Many eukaryotes are susceptible to infections by various kinds of viruses: RNA and DNA viruses, as well as retroviruses (see also Section 26.6).

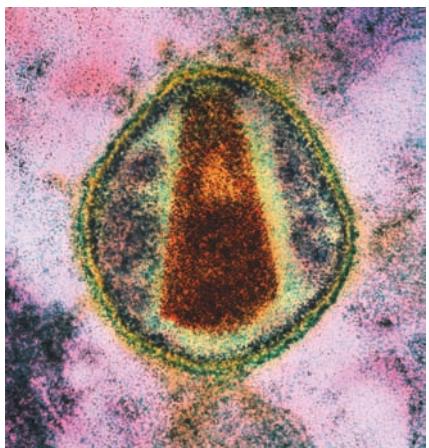
- **DNA viruses.** Many viral particles contain double-stranded DNA. However, some contain single-stranded DNA, and a complementary strand is made after the viral genome has been injected into the host cell. Like some bacteriophage, DNA viruses that infect eukaryotes are capable of undergoing both lytic and lysogenic life cycles. Examples include the herpes viruses and papillomaviruses (which cause warts).
- **RNA viruses.** Some viral genomes are made up of RNA that is usually, but not always, single-stranded. The RNA is translated by the host’s machinery to produce viral proteins, some of which are involved in replication of the RNA genome. The influenza virus has an RNA genome.



- **Retroviruses.** The retroviral genome is RNA, and the **retrovirus** encodes a protein that makes a DNA strand that is complementary to the RNA. The DNA is integrated into the host chromosome and acts as a template for both mRNA and new viral genomes. Human immunodeficiency virus (HIV) is the retrovirus that causes acquired immune deficiency syndrome (AIDS).

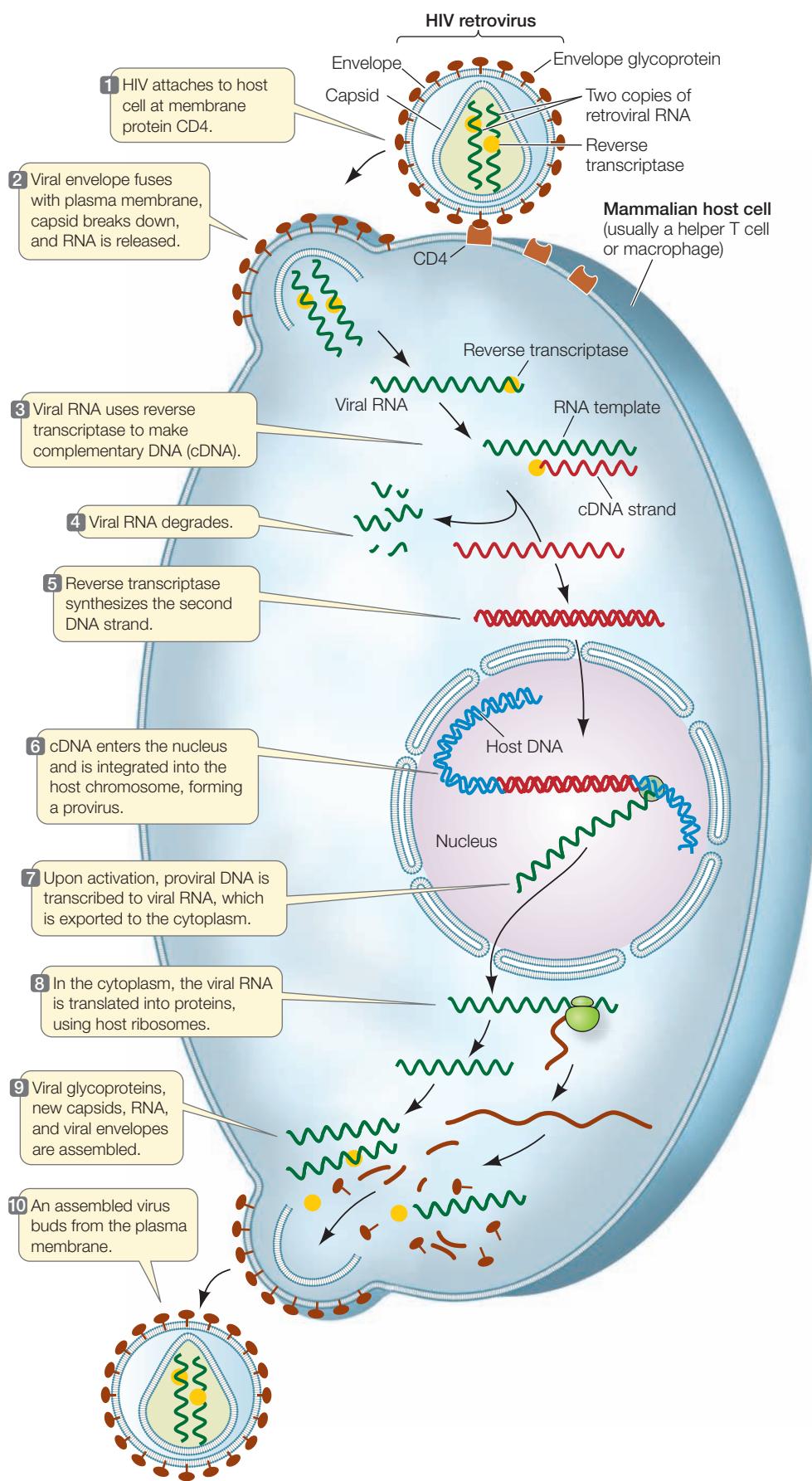
REGULATING HIV GENES As an example of viral genome regulation, we will consider the reproductive cycle of HIV (Figure 16.6). HIV is an **enveloped virus**; it is enclosed within a phospholipid membrane derived from its host cell. Proteins in the membrane are involved in infection of new host cells, which HIV enters by direct fusion of the viral envelope with the host plasma membrane.

As indicated above, a distinctive feature of the retroviral life cycle is RNA-directed DNA synthesis. This process is catalyzed by the viral enzyme **reverse transcriptase**, which uses the RNA template to produce a complementary DNA (cDNA) strand, while at the same time degrading the viral RNA. The reverse transcriptase also makes a complementary copy of the cDNA, and it is the double-stranded cDNA that gets integrated into the host’s chromosome. The integrated DNA is referred to as the **provirus** and, like the prophage, it contains promoters that are recognized by the host cell transcription apparatus. Both the reverse transcriptase and the integrase are needed for the very early stages of infection and are carried inside the HIV virion.



16.6 The Reproductive Cycle of HIV

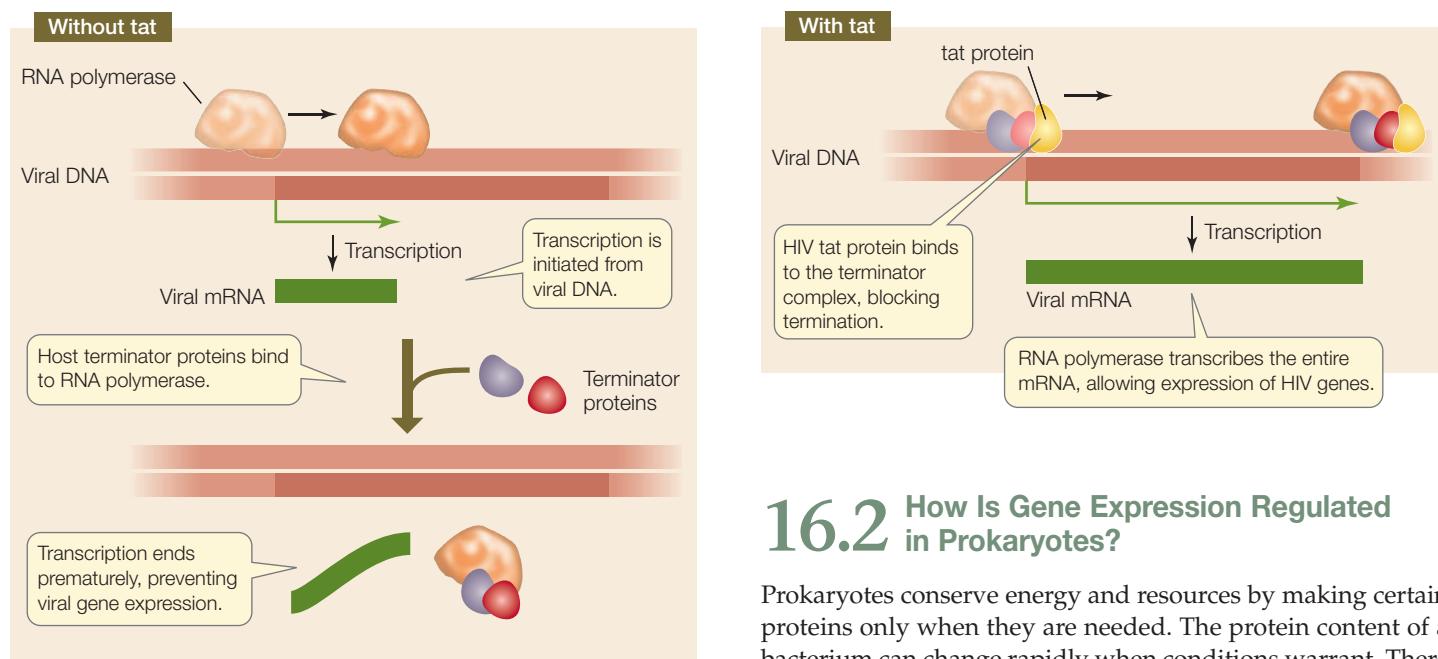
This retrovirus enters a host cell via fusion of its envelope with the host's plasma membrane. Reverse transcription of retroviral RNA then produces a DNA provirus—a molecule of complementary DNA that inserts itself into the host's genome.



The provirus resides permanently in the host chromosome and is occasionally activated to produce new virions. When this happens, the provirus is transcribed as mRNA, which is then translated by the host cell's protein-synthesizing machinery.

Under normal circumstances, the host cell regulates viral gene expression using proteins that may have originated as a defense mechanism against invaders. Host proteins bind to viral mRNA as it is being made and causes RNA polymerase to fall off the viral DNA, thereby terminating transcription. However, HIV can counteract this regulation with a virus-encoded protein called tat (*transactivator of transcription*), which binds to the terminator proteins and blocks their action. This *antitermination* allows viral gene transcription and the rest of the viral reproductive cycle to proceed (Figure 16.7).

Almost every step in the complex reproductive cycle of HIV is, in principle, a potential target for drugs to treat AIDS.



16.7 Regulation of Transcription by HIV The tat protein acts as an antiterminator, allowing transcription of the HIV genome.

16.1 RECAP

Viruses are not cells. They consist of nucleic acids and a few proteins, and require a host cell to reproduce. In the lytic cycle, the viral genome directs the host cell to generate new virions along with proteins that cause the host cell to lyse. In the lysogenic cycle, viral DNA becomes integrated in the host's genome. This DNA is multiplied along with the host cells but may remain inactive for long periods. Special viral proteins that interact with host and viral DNA sequences are the keys to the regulation of viral gene expression.

- What is the difference between positive and negative regulation of gene expression? **See Figure 16.1**
- What are the lytic and lysogenic cycles of bacteriophage? **See p. 345 and Figure 16.4**
- Describe positive and negative regulation of gene expression in bacteriophage and HIV life cycles. **See pp. 346–347 and Figures 16.5 and 16.7**

The environment surrounding prokaryotic cells can change abruptly, requiring rapid responses by the cell. We now turn to these responses, which often involve, as in viruses, the positive and negative regulation of gene expression by proteins binding to DNA.

16.2 How Is Gene Expression Regulated in Prokaryotes?

Prokaryotes conserve energy and resources by making certain proteins only when they are needed. The protein content of a bacterium can change rapidly when conditions warrant. There are several ways in which a prokaryotic cell can shut off the supply of an unneeded protein. The cell can:

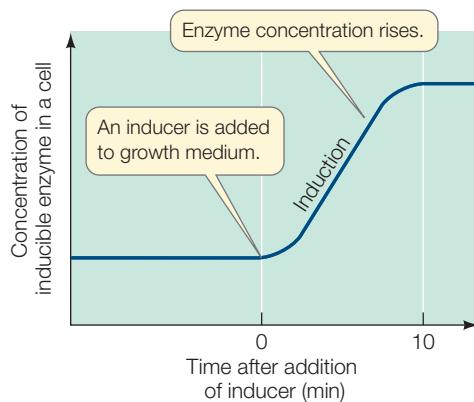
- downregulate the transcription of mRNA for that protein;
- hydrolyze the mRNA after it is made, thereby preventing translation;
- prevent translation of the mRNA at the ribosome;
- hydrolyze the protein after it is made; or
- inhibit the function of the protein.

Whichever mechanism is used, it must be both responsive to environmental signals and efficient. The earlier the cell intervenes in the process of protein synthesis, the less energy it wastes. Selective blocking of transcription is far more efficient than transcribing the gene, translating the message, and then degrading or inhibiting the protein. While all five mechanisms for regulating protein levels are found in nature, prokaryotes generally use the most efficient one: transcriptional regulation.

Regulating gene transcription conserves energy

As a normal inhabitant of the human intestine, *E. coli* must be able to adjust to sudden changes in its chemical environment. Its host may present it with one foodstuff one hour (e.g., glucose) and another the next (e.g., lactose). Such changes in nutrients present the bacterium with a metabolic challenge. Glucose is its preferred energy source, and is the easiest sugar to metabolize, but not all of its host's foods contain an abundant supply of glucose. For example, the bacterium may suddenly be deluged with milk, whose main sugar is lactose. Lactose is a β -galactoside—a disaccharide containing galactose β -linked to glucose (see Section 3.3). Three proteins are involved in the initial uptake and metabolism of lactose by *E. coli*:

- β -galactoside permease is a carrier protein in the bacterial plasma membrane that moves the sugar into the cell.



- β -galactosidase is an enzyme that hydrolyses lactose to glucose and galactose.
- β -galactoside transacetylase transfers acetyl groups from acetyl CoA to certain β -galactosides. Its role in the metabolism of lactose is not clear.

When *E. coli* is grown on a medium that contains glucose but no lactose or other β -galactosides, the levels of these three proteins are extremely low—the cell does not waste energy and materials making the unneeded enzymes. But if the environment changes such that lactose is the predominant sugar available and very little glucose is present, the bacterium promptly begins making all three enzymes. There are only two molecules of β -galactosidase present in an *E. coli* cell when glucose is present in the medium. But when glucose is absent, the presence of lactose can induce the synthesis of 3,000 molecules of β -galactosidase per cell!

If lactose is removed from *E. coli*'s environment, synthesis of the three enzymes stops almost immediately. The enzyme molecules already present do not disappear; they are merely diluted during subsequent cell divisions until their concentration falls to the original low level within each bacterium.

Compounds that, like lactose, stimulate the synthesis of a protein are called **inducers** (Figure 16.8). The proteins that are

16.8 An Inducer Stimulates the Expression of a Gene for an Enzyme It is most efficient for a cell to produce an enzyme only when it is needed. Some enzymes are induced by the presence of the substance they act upon (for example, β -galactosidase is induced by the presence of lactose).

produced are called **inducible proteins**, whereas proteins that are made all the time at a constant rate are called **constitutive proteins**. (Think of the constitution of a country, a document that does not change under normal circumstances.)

We have now seen two basic ways of regulating the rate of a metabolic pathway. In Section 8.5 we described allosteric regulation of enzyme activity (the rate of enzyme-catalyzed reactions); this mechanism allows rapid fine-tuning of metabolism. Regulation of protein synthesis—that is, regulation of the concentration of enzymes—is slower, but results in greater savings of energy and resources. Protein synthesis is a highly endergonic process, since assembling mRNA, charging tRNA, and moving the ribosomes along mRNA all require the hydrolysis of ATP. Figure 16.9 compares these two modes of regulation.

Operons are units of transcriptional regulation in prokaryotes

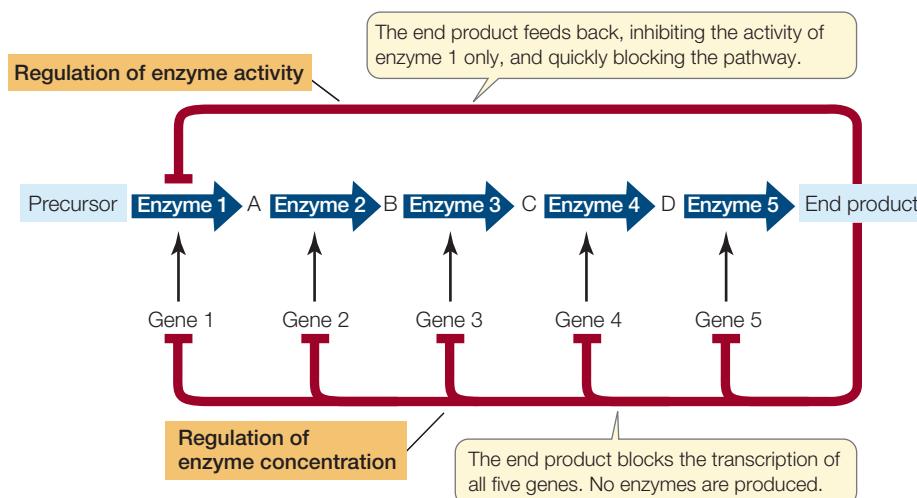
The genes that encode the three enzymes for processing lactose in *E. coli* are **structural genes**; they specify the primary structure (the amino acid sequence) of a protein molecule. Structural genes are genes that can be transcribed into mRNA.

The three structural genes involved in the metabolism of lactose lie adjacent to one another on the *E. coli* chromosome. This arrangement is no coincidence: the genes share a single promoter, and their DNA is transcribed into a single, continuous molecule of mRNA. Because this particular mRNA governs the synthesis of all three lactose-metabolizing enzymes, either all or none of these enzymes are made, depending on whether their common message—their mRNA—is present in the cell.

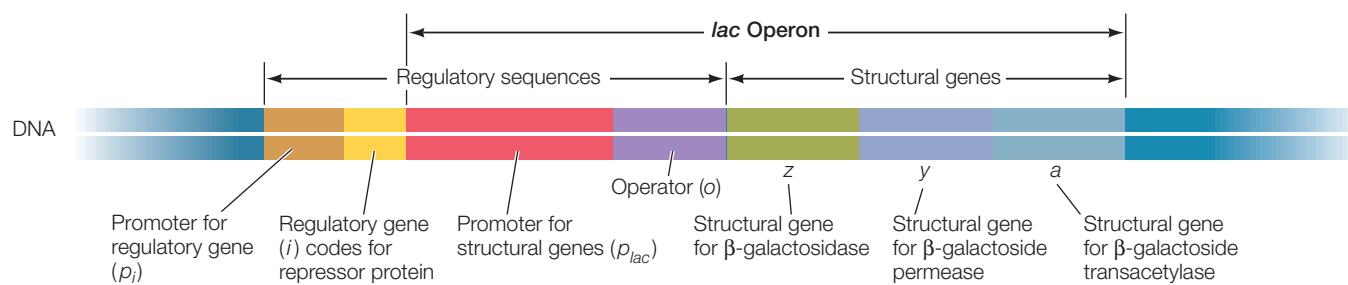
A cluster of genes with a single promoter is called an **operon**, and the operon that encodes the three lactose-metabolizing

enzymes in *E. coli* is called the *lac operon*. The *lac* operon promoter can be very efficient (the maximum rate of mRNA synthesis can be high) but mRNA synthesis can be shut down when the enzymes are not needed. This example of negative regulation was elegantly worked out by Nobel Prize winners François Jacob and Jacques Monod.

In addition to the promoter, an operon has other regulatory sequences that are not transcribed. A typical operon consists of a promoter, an operator, and two or more



16.9 Two Ways to Regulate a Metabolic Pathway Feedback from the end product of a metabolic pathway can block enzyme activity (allosteric regulation), or it can stop the transcription of genes that code for the enzymes in the pathway (transcriptional regulation).



16.10 The *lac* Operon of *E. coli* The *lac* operon of *E. coli* is a segment of DNA that includes a promoter, an operator, and the three structural genes that code for lactose-metabolizing enzymes.

structural genes (Figure 16.10). The **operator** is a short stretch of DNA that lies between the promoter and the structural genes. It can bind very tightly with regulatory proteins that either activate or repress transcription. There are numerous mechanisms to control the transcription of operons; here we will focus on three examples:

- An inducible operon regulated by a repressor protein
- A repressible operon regulated by a repressor protein
- An operon regulated by an activator protein

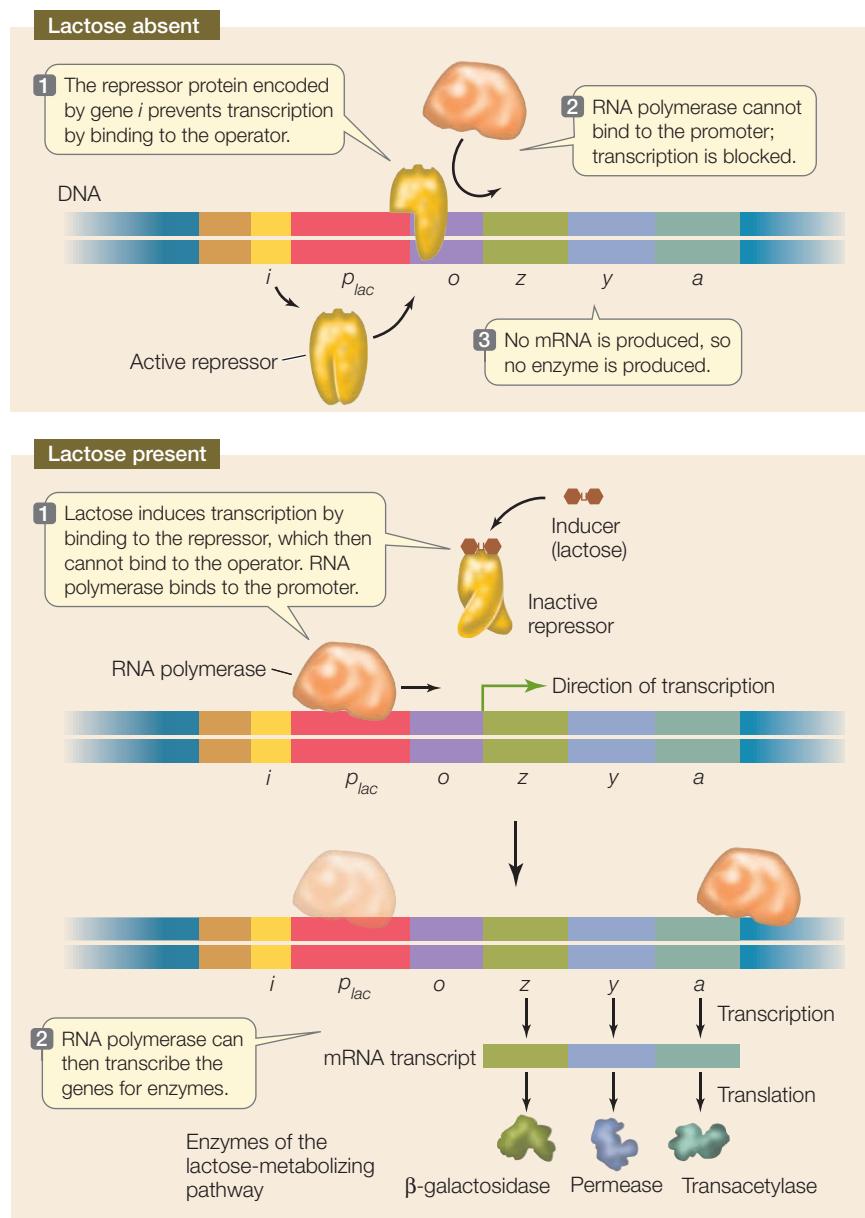
Operator-repressor interactions control transcription in the *lac* and *trp* operons

The *lac* operon contains a promoter, to which RNA polymerase binds to initiate transcription, and an operator, to which a **repressor** protein can bind. When the repressor is bound, transcription of the operon is blocked.

The repressor protein has two binding sites: one for the operator and the other for the inducer, lactose. Binding with the inducer changes the shape of the repressor protein. This change in three-dimensional structure (conformation) prevents the repressor from binding to the operator (Figure 16.11). As a result, RNA polymerase can bind to the promoter and start transcribing the structural genes of the *lac* operon.

Study Figure 16.11 for the features of this negative control. You will notice that:

- in the absence of inducer, the operon is turned off;
- control is exerted by a regulatory protein—the repressor—that turns the operon off;
- the inducer, when present, binds to and changes the shape of the repressor so that it no longer binds to the operator, turning the operon on;

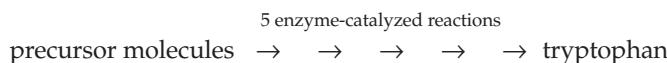


16.11 The *lac* Operon: An Inducible System Lactose (the inducer) leads to synthesis of the enzymes in the lactose-metabolizing pathway by binding to the repressor protein and preventing its binding to the operator.

- the **regulatory gene** produces a protein whose sole function is to regulate expression of the other genes; and
- certain DNA sequences (operators and promoters) do not code for proteins, but are binding sites for regulatory or other proteins.

In contrast to the inducible system of the *lac* operon, other operons in *E. coli* are repressible; that is, they are repressed when molecules called **co-repressors** bind to their repressors. This binding causes the repressor to change shape and bind to the operator, thereby inhibiting transcription.

An example is the operon whose structural genes catalyze the synthesis of the amino acid tryptophan:



When tryptophan is present in the cell in adequate concentrations, it is advantageous to stop making the enzymes for tryptophan synthesis. To do this, the cell uses a repressor that binds to an operator upstream of the genes of the *trp* operon. But the repressor of the *trp* operon is not normally bound to the operator; it only binds when its shape is changed by binding to tryptophan, the co-repressor. To summarize the differences between these two types of operons:

- In *inducible* systems, the substrate of a metabolic pathway (the inducer) interacts with a regulatory protein (the repressor), rendering the repressor incapable of binding to the operator and thus allowing transcription.
- In *repressible* systems, the product of a metabolic pathway (the co-repressor) binds to a regulatory protein, which is then able to bind to the operator and block transcription.

yourBioPortal.com

GO TO Animated Tutorial 16.2 • The *trp* Operon

In general, inducible systems control catabolic pathways (which are turned on only when the substrate is available), whereas repressible systems control anabolic pathways (which are turned on until the concentration of the product becomes excessive). In both of the systems described here, the regulatory protein is a repressor that functions by binding to the operator. Next we will consider an example of positive control involving an activator.

16.12 Catabolite Repression Regulates the *lac* Operon The promoter for the *lac* operon does not function efficiently in the absence of cAMP, as occurs when glucose levels are high. High glucose levels thus repress the enzymes that metabolize lactose.

Protein synthesis can be controlled by increasing promoter efficiency

The examples described in the previous section are termed negative control because transcription is *decreased* in the presence of a repressor protein. *E. coli* can also use positive control to *increase* transcription through the presence of an **activator** protein. For an example we return to the *lac* operon, where the relative levels of glucose and lactose determine the amount of transcription. When lactose is present and glucose is low, the *lac* operon is activated by binding of a protein called cAMP receptor protein (CRP) to the *lac* operon promoter. CRP is an activator of transcription, because its binding results in more efficient binding of RNA polymerase to the promoter, and thus increased transcription of the structural genes (**Figure 16.12**).

In the presence of abundant glucose, CRP does not bind to the promoter and so the efficiency of transcription of the *lac* operon is reduced. This is an example of **catabolite repression**, a system of gene regulation in which the presence of the preferred energy source represses other catabolic pathways. The signaling pathway that controls catabolite repression of the *lac* operon involves the second messenger cAMP (see Section 7.3). The mechanisms controlling positive and negative regulation of the *lac* operon are summarized in **Table 16.1**.

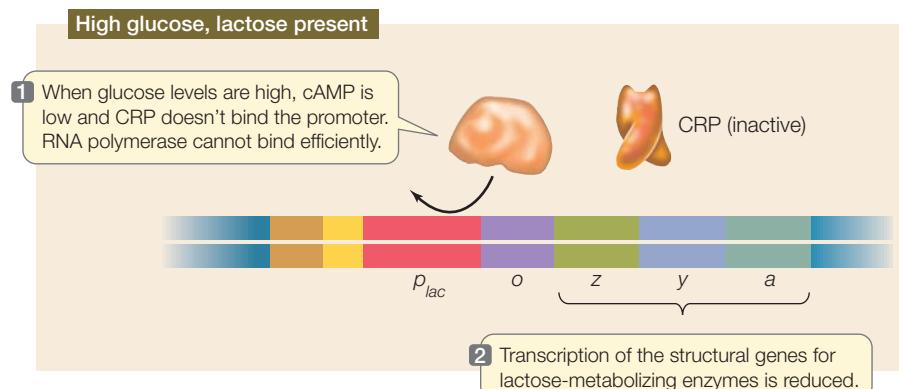
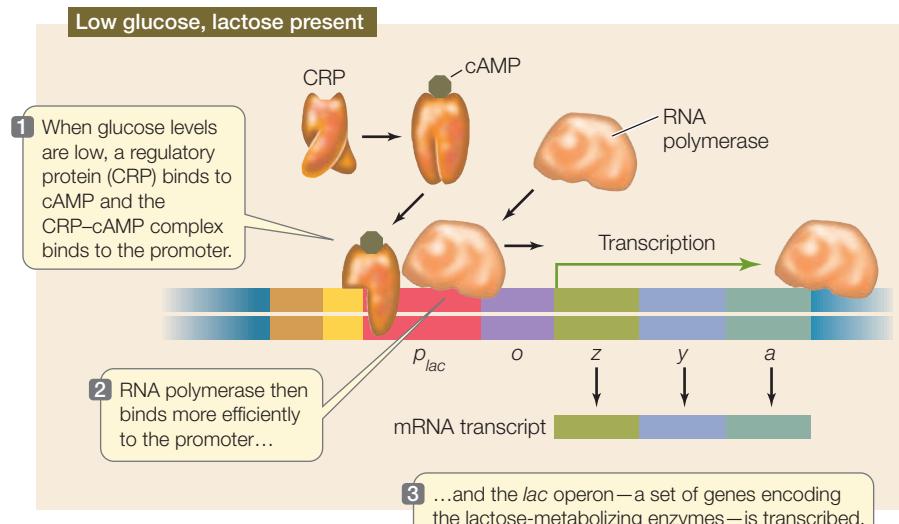


TABLE 16.1
Positive and Negative Regulation in the *lac* Operon^a

GLUCOSE	cAMP LEVELS	RNA POLYMERASE BINDING TO PROMOTER	LACTOSE	LAC REPRESSOR	TRANSCRIPTION OF <i>lac</i> GENES?	LACTOSE USED BY CELLS?
Present	Low	Absent	Absent	Active and bound to operator	No	No
Present	Low	Present, not efficient	Present	Inactive and not bound to operator	Low level	No
Absent	High	Present, very efficient	Present	Inactive and not bound to operator	High level	Yes
Absent	High	Absent	Absent	Active and bound to operator	No	No

^aNegative regulators are in red type.

16.2 RECAP

Gene expression in prokaryotes is most commonly regulated through control of transcription. An operon consists of a set of closely linked structural genes and the DNA sequences (promoter and operator) that control their transcription. Operons can be regulated by both negative and positive controls.

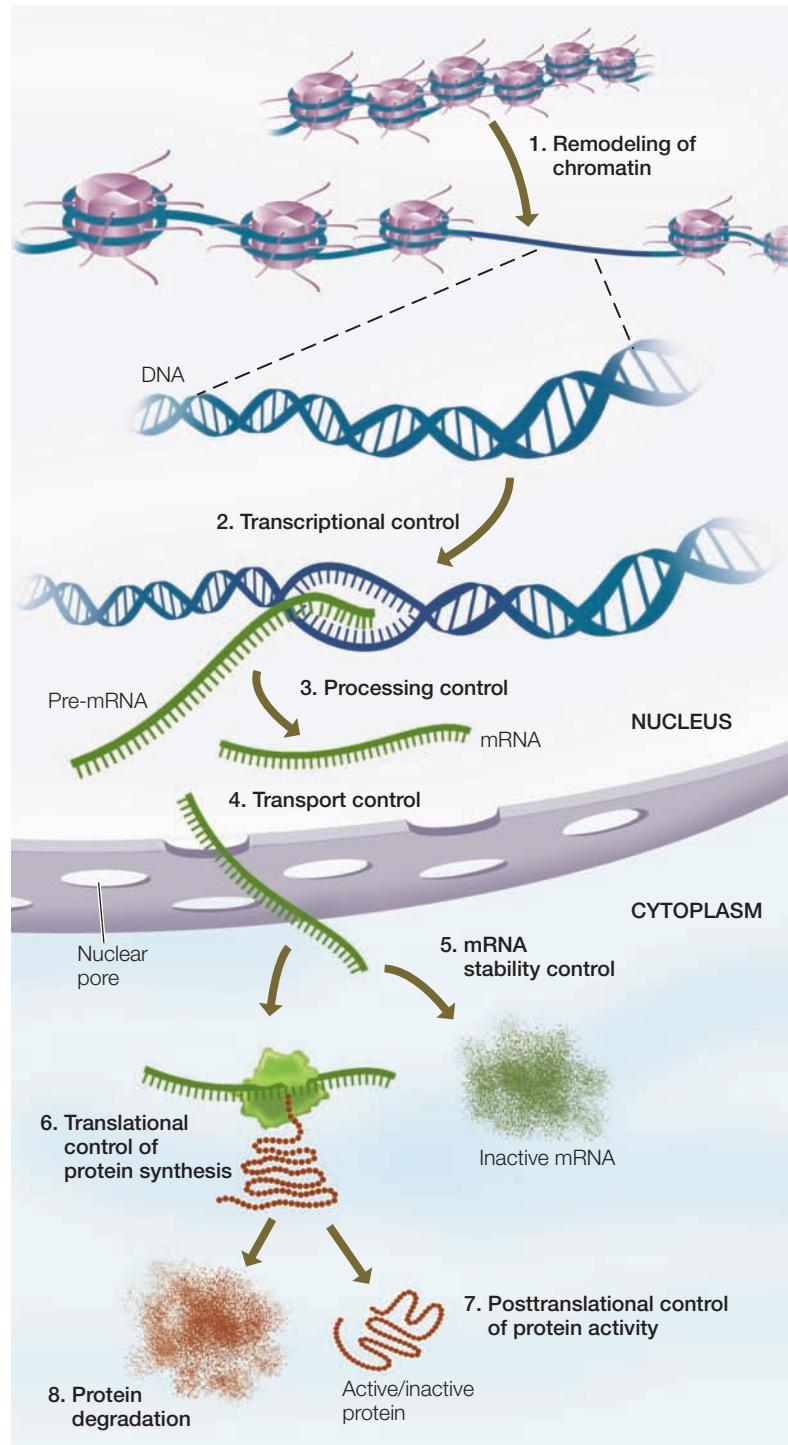
- Describe the molecular conditions at the *lac* operon promoter in the presence versus absence of lactose. See Figure 16.11
- What are the key differences between an inducible system and a repressible system? See p. 351
- What are the differences between positive and negative control of transcription? See p. 351 and Table 16.1

Studies of viruses and bacteria provide a basic understanding of mechanisms that regulate gene expression and of the roles of regulatory proteins in both positive and negative regulation. We now turn to the control of gene expression in eukaryotes. You will see both negative and positive control of transcription, as well as posttranscriptional mechanisms of regulation.

16.3 How Is Eukaryotic Gene Transcription Regulated?

For the normal development of an organism from fertilized egg to adult, and for each cell to acquire and maintain its proper specialized function, certain proteins must be made at just the right times and in just the right cells; these proteins must not be made at other times in other cells. Thus the expression of eukaryotic genes must be precisely regulated.

As in prokaryotes, eukaryotic gene expression can be regulated at a number of different points in the process of transcribing and translating the gene into a protein (Figure 16.13). In this section we will describe the mechanisms



16.13 Potential Points for the Regulation of Gene Expression

Gene expression can be regulated before transcription (1), during transcription (2, 3), after transcription but before translation (4, 5), at translation (6), or after translation (7).

TABLE 16.2
Transcription in Prokaryotes and Eukaryotes

PROKARYOTES	EUKARYOTES
Locations of functionally related genes	Often clustered in operons
RNA polymerases	One
Promoters and other regulatory sequences	Few
Initiation of transcription	Binding of RNA polymerase to promoter
	Binding of many proteins, including RNA polymerase

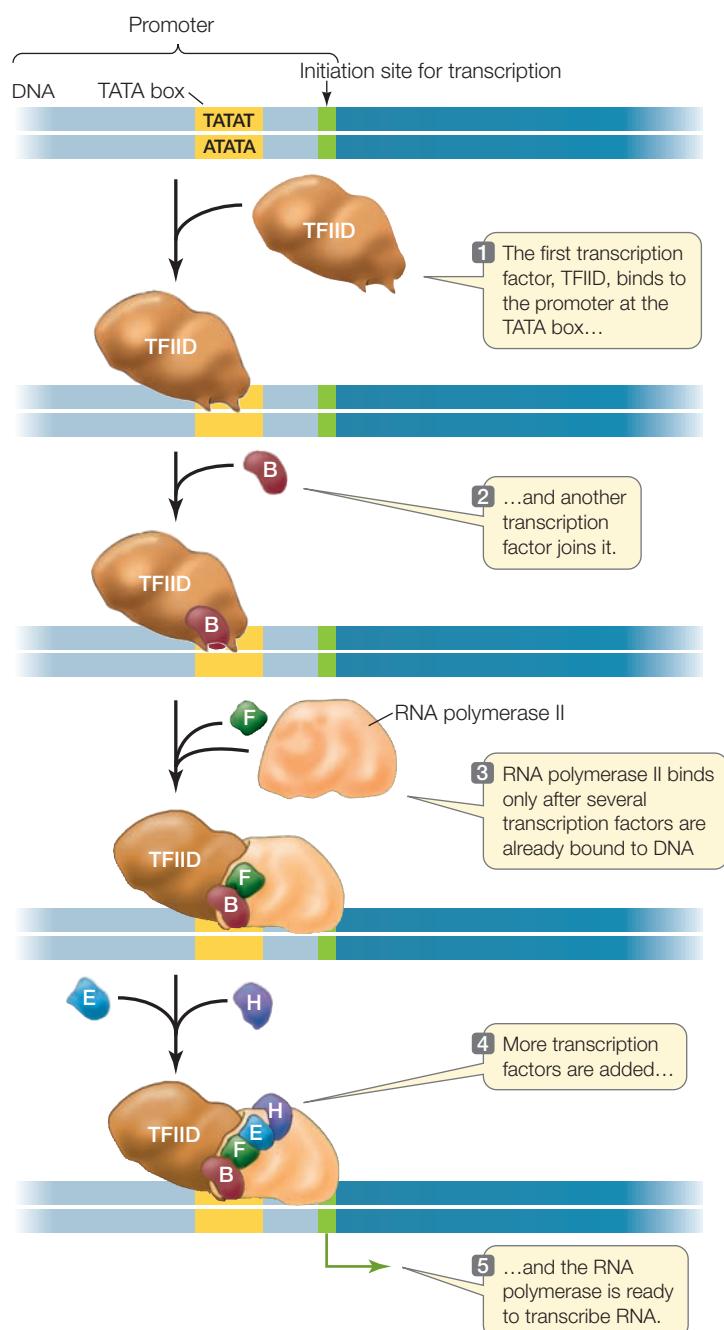
that result in the selective transcription of specific genes. The mechanisms for regulating gene expression in eukaryotes have similar themes to those of prokaryotes. Both types of cells use DNA–protein interactions and negative and positive control. However, there are many differences, some of them dictated by the presence of a nucleus, which physically separates transcription and translation (**Table 16.2**).

Transcription factors act at eukaryotic promoters

As in prokaryotes, a promoter in eukaryotes is a sequence of DNA near the 5' end of the coding region of a gene where RNA polymerase binds and initiates transcription. There are typically two important sequences in a promoter: One is the **recognition sequence**—the sequence recognized by RNA polymerase. The second, closer to the transcription initiation site, is the **TATA box** (so called because it is rich in AT base pairs), where DNA begins to denature so that the template strand can be exposed.

Eukaryotic RNA polymerase II cannot simply bind to the promoter and initiate transcription. Rather, it does so only after various regulatory proteins, called **transcription factors**, have assembled on the chromosome (**Figure 16.14**). First, the protein TFIID (“TF” stands for transcription factor) binds to the TATA box. Binding of TFIID changes both its own shape and that of the DNA, presenting a new surface that attracts the binding of other transcription factors to form a transcription complex. RNA polymerase II binds only after several other proteins have bound to this complex.

Some regulatory DNA sequences, such as the TATA box, are common to the promoters of many eukaryotic genes and are recognized by transcription factors that are found in all the cells of an organism. Other sequences found in promoters are specific



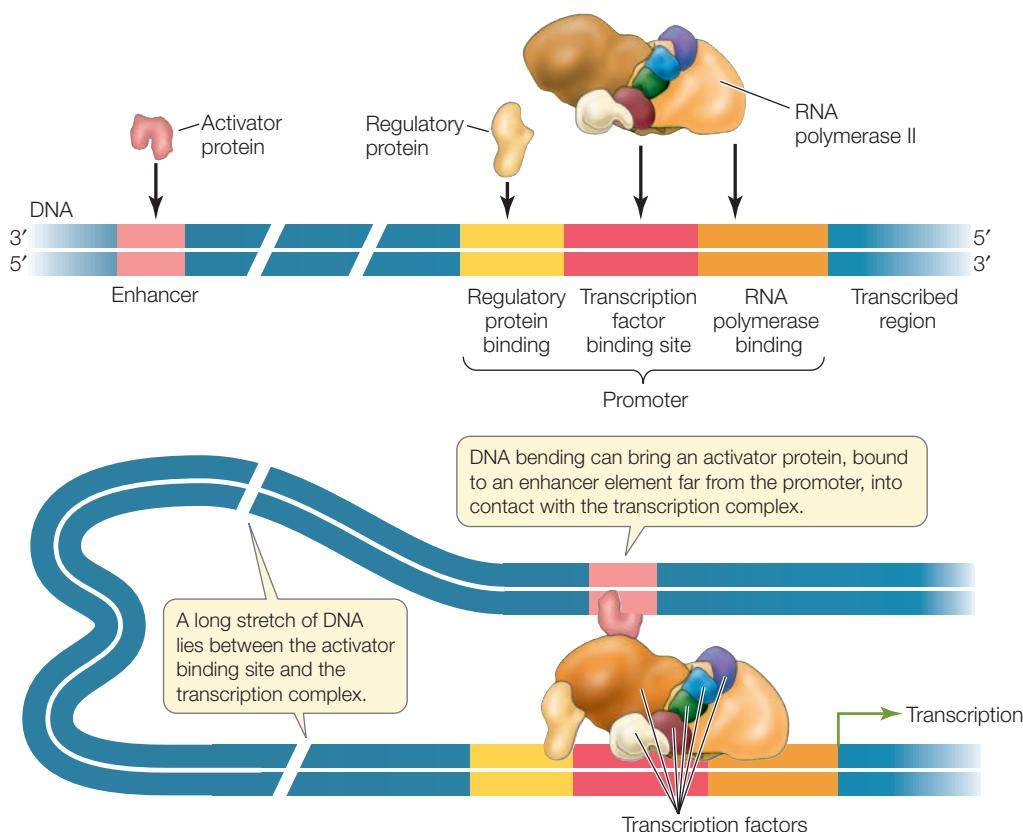
16.14 The Initiation of Transcription in Eukaryotes Apart from TFIID, which binds to the TATA box, each transcription factor in this transcription complex has binding sites only for the other proteins in the complex, and does not bind directly to DNA. B, E, F, and H are transcription factors.

to only a few genes and are recognized by transcription factors found only in certain types of cells. These specific transcription factors play an important role in cell differentiation, the structural and functional specialization of cells during development.

Other proteins can recognize and bind to DNA sequences and regulate transcription

In addition to the promoter, there are other short sequences (elements) of DNA that bind regulatory proteins, which in turn interact with RNA polymerase to regulate the rate of transcription (Figure 16.15). Some of these DNA elements are positive regulators (termed enhancers, which bind activator proteins) and others are negative (silencers, which bind repressor proteins). Some occur near the promoter and others as far as 20,000 base pairs away. One example of a transcription factor is CREB, which you read about in the opening essay of this chapter. When the activators and/or repressors (collectively termed transcription factors) bind to these elements, they interact with the RNA polymerase complex, causing DNA to bend. Often many such binding proteins are involved, and *the combination of factors present determines the rate of transcription*.

For example, the immature red blood cells in bone marrow make large amounts of β -globin. At least thirteen different transcription factors are involved in regulating transcription of the β -globin gene in these cells. Not all of these factors are present or active in other cells, such as the immature white blood cells produced by the same bone marrow. As a result the β -globin gene is not transcribed in those cells. So although the same



genes are present in all cells, the fate of the cell is determined by which of its genes are expressed. How do transcription factors recognize specific DNA sequences?

Specific protein-DNA interactions underlie binding

As we have seen, transcription factors with specific DNA binding domains are involved in the activation and inactivation of specific genes. There are four common structural themes in the protein domains that bind to DNA. These themes, or **structural motifs**, consist of different combinations of structural elements (protein conformations) and may include special components such as zinc. The four common structural motifs in DNA binding domains are: helix-turn-helix, leucine zipper, zinc finger, and helix-loop-helix (Figure 16.16).

Let's look at how one of these motifs works. As pointed out in Section 13.2, the complementary bases in DNA not only form hydrogen bonds with each other, but also can form additional hydrogen bonds with proteins, particularly at points exposed in the major and minor grooves. In this way, an intact DNA double helix can be recognized by a protein motif whose structure:

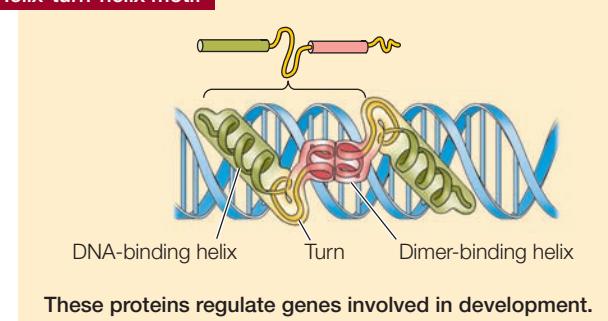
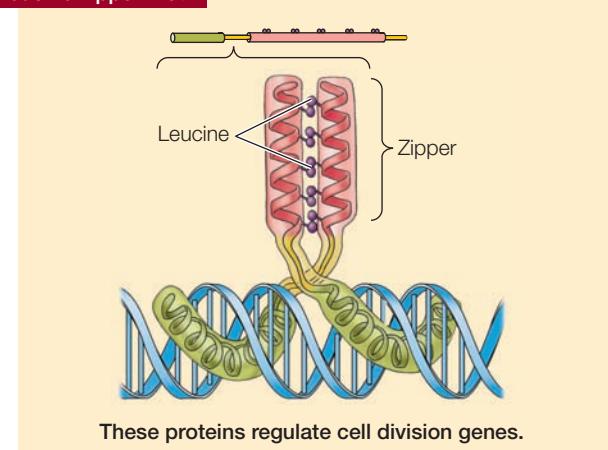
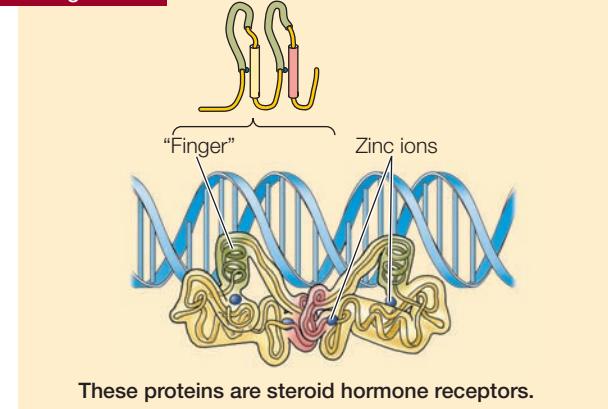
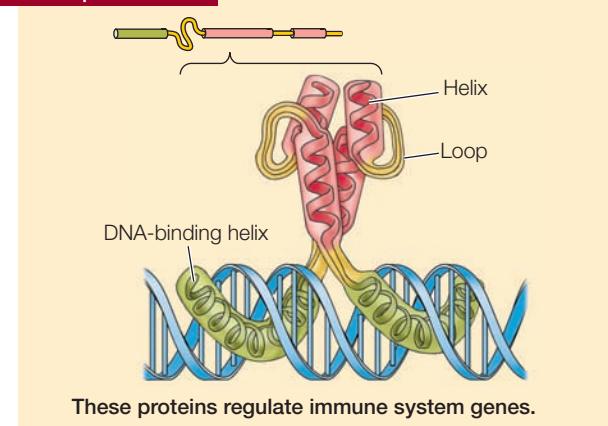
- fits into the major or minor groove;
- has amino acids that can project into the interior of the double helix; and
- has amino acids that can form hydrogen bonds with the interior bases.

The helix-turn-helix motif, in which two α -helices are connected via a non-helical turn, fits these three criteria. The interior-facing "recognition" helix is the one whose amino acids interact with the bases inside the DNA. The exterior-facing he-

lix sits on the sugar-phosphate backbone, ensuring that the interior helix is presented to the bases in the correct configuration. Many repressor proteins have this helix-turn-helix motif in their structure.

Repressors can inhibit transcription in several different ways. They can prevent the binding of transcriptional activators to DNA, or they can interact with other DNA binding proteins to decrease the rate of transcription.

16.15 Transcription Factors, Repressors, and Activators The actions of many proteins determine whether and where RNA polymerase II will transcribe DNA.

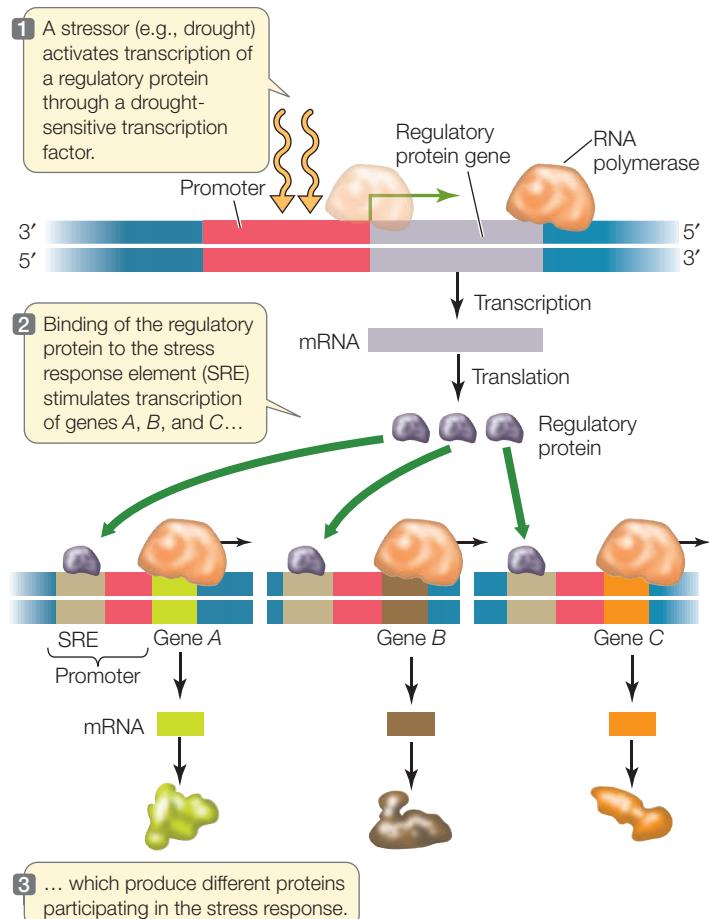
Helix-turn-helix motif**Leucine zipper motif****Zinc finger motif****Helix-loop-helix motif**

16.16 Protein–DNA Interactions The DNA-binding domains of most regulatory proteins contain one of four structural motifs.

The expression of sets of genes can be coordinately regulated by transcription factors

How do eukaryotic cells coordinate the regulation of several genes whose transcription must be turned on at the same time? Prokaryotes solve this problem by arranging multiple genes in an operon that is controlled by a single promoter. But most eukaryotic genes have their own separate promoters, and genes that are coordinately regulated may be far apart. In these cases, the expression of genes can be coordinated if they share regulatory sequences that bind the same transcription factors.

This type of coordination is used by organisms to respond to stress—for example, by plants in response to drought. Under conditions of drought stress, a plant must simultaneously synthesize a number of proteins whose genes are scattered throughout the genome. The synthesis of these proteins comprises the stress response. To coordinate expression, each of these genes has a specific regulatory sequence near its promoter called the *stress response element (SRE)*. A transcription factor binds to this element and stimulates mRNA synthesis (Figure 16.17). The stress re-



16.17 Coordinating Gene Expression A single environmental signal, such as drought stress, causes the synthesis of a transcriptional regulatory protein that acts on many genes.

sponse proteins not only help the plant conserve water, but also protect the plant against excess salt in the soil and freezing. This finding has considerable importance for agriculture because crops are often grown under less than optimal conditions.

16.3 RECAP

A number of transcription factors must bind to a eukaryotic promoter before RNA polymerase will bind to it and begin transcription. This provides a number of ways to increase or decrease transcription.

- Describe some of the different ways in which transcription factors regulate gene transcription. See pp. 353–354 and Figure 16.15
- How can more than one gene be regulated at the same time? See p. 355 and Figure 16.17

The mechanisms for control of gene expression that we have discussed so far involve direct interactions between proteins and specific DNA elements. If the sequences of the DNA elements are altered, then transcription of the gene will be affected. However, there are other mechanisms for controlling gene expression that do not depend on specific DNA sequences. We will discuss these mechanisms in the next section.

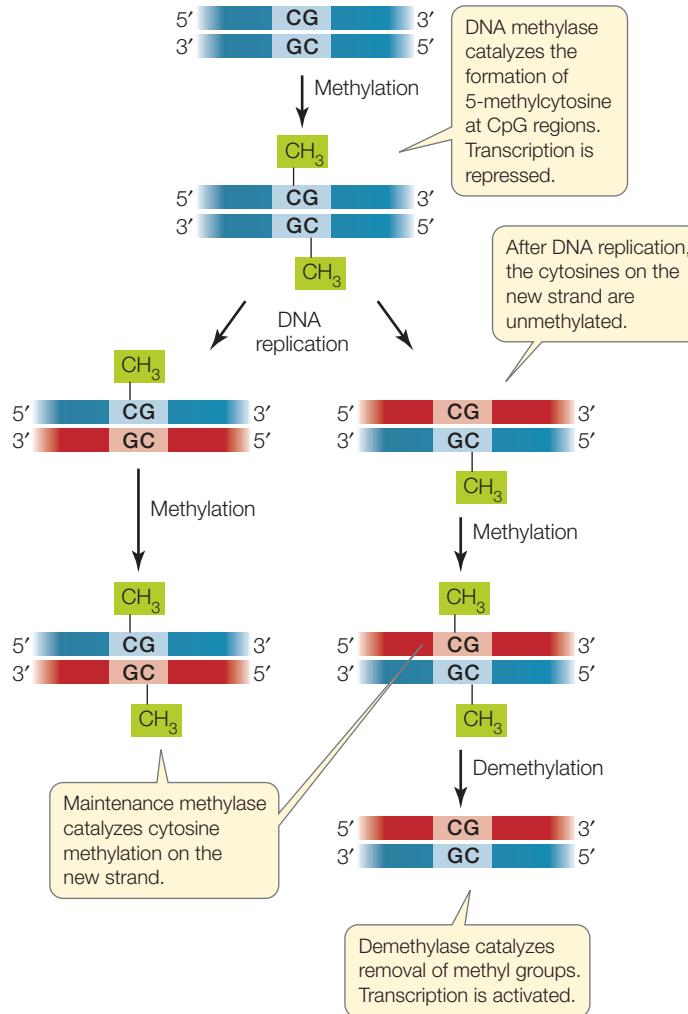
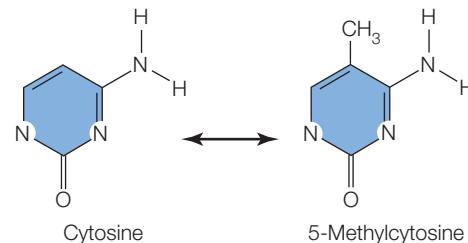
16.4 How Do Epigenetic Changes Regulate Gene Expression?

In the mid-twentieth century, the great developmental biologist Conrad Hal Waddington coined the term “epigenetics” and defined it as “that branch of biology which studies the causal interactions between genes and their products which bring the phenotype into being.” Today **epigenetics** is defined more specifically, referring to changes in the expression of a gene or set of genes that occur without changing the DNA sequence. These changes are reversible, but sometimes are stable and heritable. They include two processes: **DNA methylation** and chromosomal protein alterations.

DNA methylation occurs at promoters and silences transcription

Depending on the organism, from 1 to 5 percent of cytosine residues in the DNA are chemically modified by the addition of a methyl group ($-\text{CH}_3$) to the 5'-carbon, to form 5-methylcytosine (Figure 16.18). This covalent addition is catalyzed by the enzyme **DNA methyltransferase** and, in mammals, usually occurs in C residues that are adjacent to G residues. DNA regions rich in these doublets are called **CpG islands**, and are especially abundant in promoters.

This covalent change in DNA is heritable: when DNA is replicated, a **maintenance methylase** catalyzes the formation of 5-methylcytosine in the new DNA strand. However, the pattern of cytosine methylation can also be altered, because methylation is reversible: a third enzyme, appropriately called **demethylase**,



16.18 DNA Methylation: an Epigenetic Change The reversible formation of 5-methylcytosine in DNA can alter the rate of transcription.

lase, catalyzes the removal of the methyl group from cytosine (see Figure 16.18).

What is the effect of DNA methylation? During replication and transcription, 5-methylcytosine behaves just like plain cytosine: it base pairs with guanine. But extra methyl groups in a promoter attract proteins that bind methylated DNA. These proteins are generally involved in the repression of gene transcription; thus heavily methylated genes tend to be inactive. This form of genetic regulation is epigenetic because it affects gene expression patterns without altering the DNA sequence.

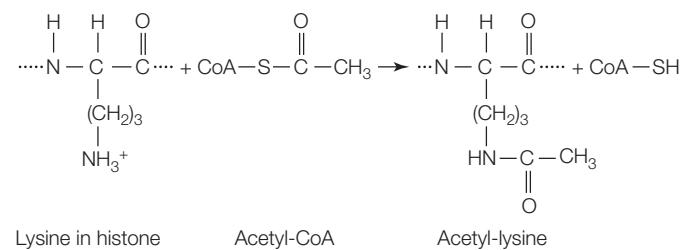
DNA methylation is important in development from egg to embryo. For example, when a mammalian sperm enters an egg, many genes in first the male and then the female genome become demethylated. Thus many genes that are usually inactive are expressed during early development. As the embryo develops and its cells become more specialized, genes whose products are not needed in particular cell types become methylated. These methylated genes are “silenced”; their transcription is repressed. However, unusual or abnormal events can sometimes turn silent genes back on.

For example, DNA methylation may play roles in the genesis of some cancers. In cancer cells, oncogenes get activated and promote cell division, and tumor suppressor genes (that normally inhibit cell division) are turned off (see Chapter 11). This misregulation can occur when the promoters of oncogenes become demethylated while those of tumor suppressor genes become methylated. This is the case in colorectal cancer.

Histone protein modifications affect transcription

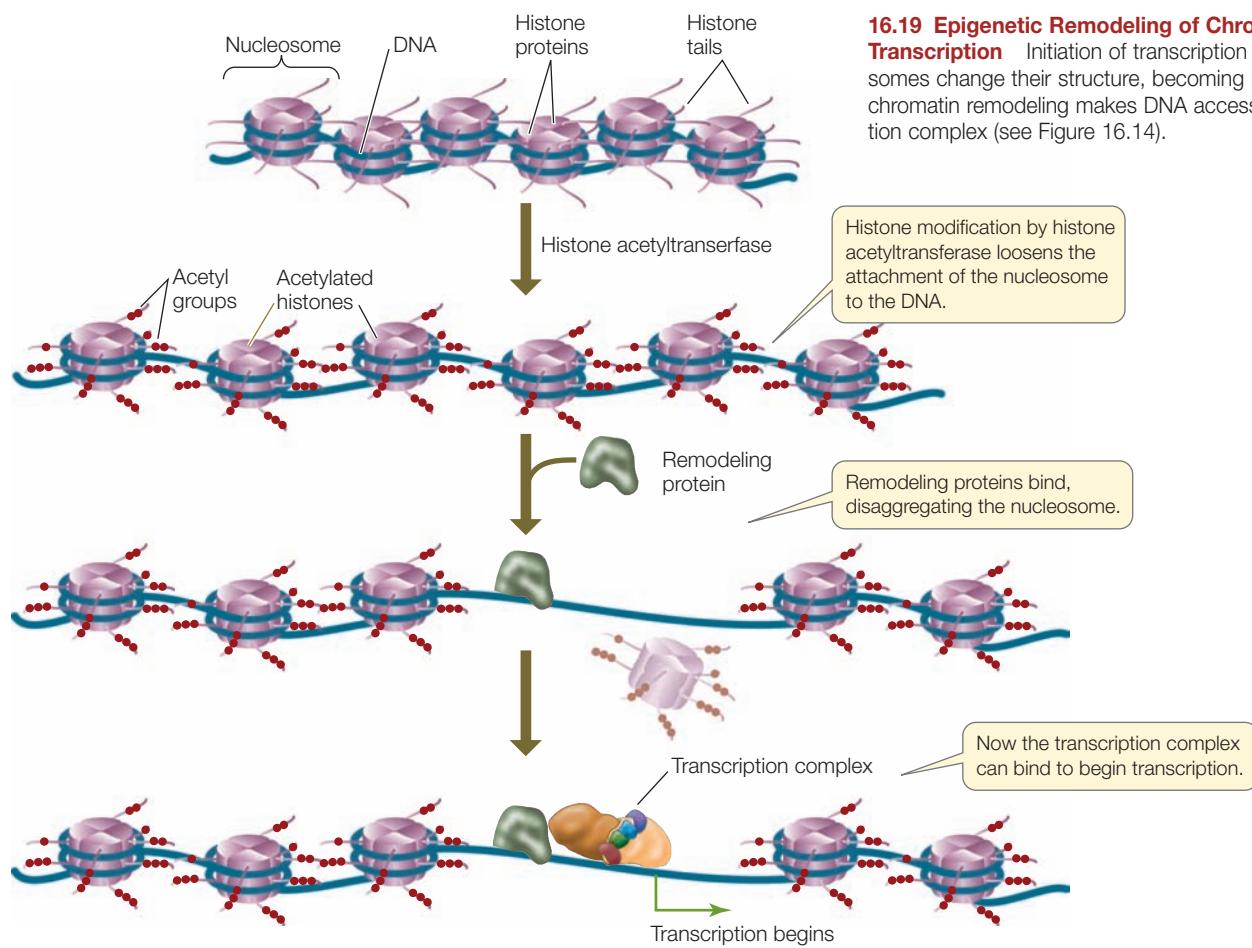
Another mechanism for epigenetic gene regulation is the alteration of chromatin structure, or *chromatin remodeling*. DNA is packaged with histone proteins into nucleosomes, which can make DNA physically inaccessible to RNA polymerase and the rest of the transcription apparatus. Each histone protein has a

“tail” of approximately 20 amino acids at its N terminus that sticks out of the compact structure and contains certain positively charged amino acids (notably lysine). Enzymes called histone acetyltransferases can add acetyl groups to these positively charged amino acids, thus changing their charges:



Ordinarily, there is strong electrostatic attraction between the positively charged histone proteins and DNA, which is negatively charged because of its phosphate groups. Reducing the positive charges of the histone tails reduces the affinity of the histones for DNA, opening up the compact nucleosome. Additional chromatin remodeling proteins can bind to the loosened nucleosome–DNA complex, opening up the DNA for gene expression (**Figure 16.19**). Histone acetyltransferases can thus activate transcription.

Another kind of chromatin remodeling protein, histone deacetylase, can remove the acetyl groups from histones and



thereby repress transcription. Histone deacetylases are targets for drug development to treat some forms of cancer. As noted above, certain genes block cell division in normal specialized tissues. In some cancers these genes are less active than normal, and the histones near them show excessive levels of deacetylation. Theoretically, a drug acting as a histone deacetylase inhibitor could tip the balance toward acetylation and this might activate genes that normally inhibit cell division.

Other types of histone modification can affect gene activation and repression. For example, histone methylation is associated with gene inactivation and histone phosphorylation also affects gene expression, the specific effect depending on which amino acid is modified. All of these effects are reversible and so the activity of a eukaryotic gene may be determined by very complex patterns of histone modification. David Allis of the Rockefeller University in New York City has dubbed this epigenetic system the “histone code.”

Epigenetic changes induced by the environment can be inherited

Despite the fact that they are reversible, many epigenetic changes such as DNA methylation and histone modification can permanently alter gene expression patterns in a cell. If the cell is a germ line cell that forms gametes, the epigenetic changes can be passed on to the next generation. But what determines these epigenetic changes? A clue comes from a recent study of monozygotic twins.

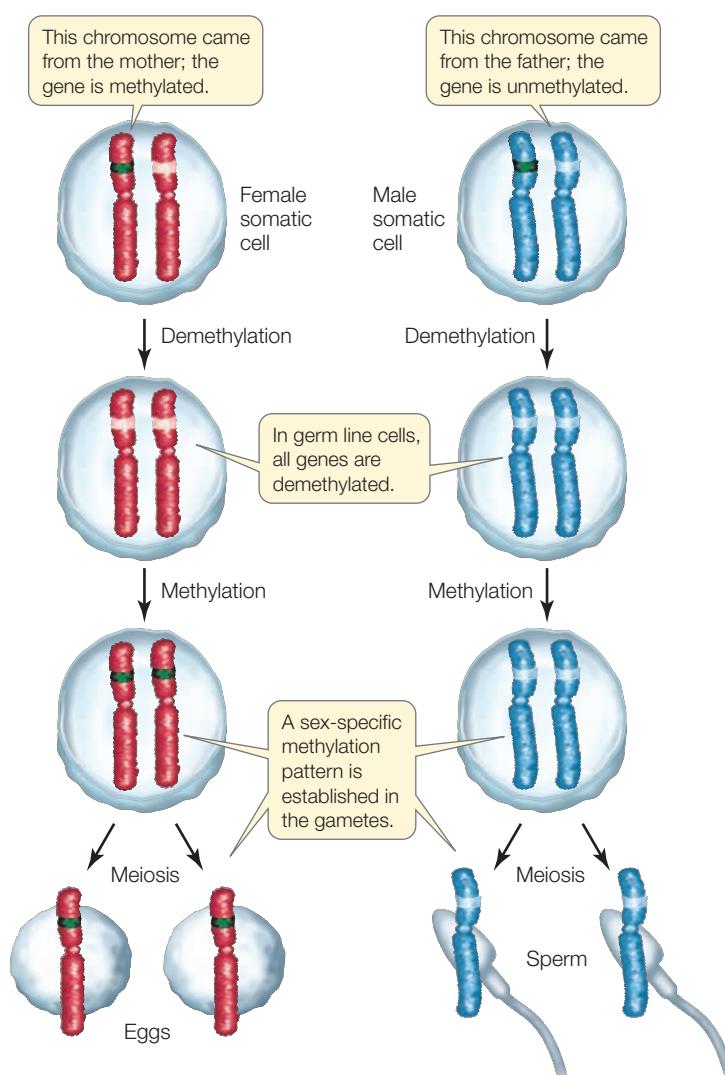
Monozygotic twins come from a single fertilized egg that divides to produce two separate cells; each of these goes on to develop a separate individual. Twin brothers or sisters thus have identical genomes. But are they identical in their *epigenomes*? A comparison of DNA in hundreds of such twin pairs shows that in tissues of three-year-olds, the DNA methylation patterns are virtually the same. But by age 50, by which time the twins have usually been living apart for decades, in different environments, the patterns are quite different. This indicates that the *environment plays an important role in epigenetic modifications* and, therefore, in the regulation of genes that these modifications affect.

What factors in the environment lead to epigenetic changes? One might be stress: when mice are put in a stressful situation, genes that are involved in important brain pathways become heavily methylated (and transcriptionally inactive). Treatment of the stressed mice with an antidepressant drug “hits the undo button,” reversing these changes. Transcription factors such as CREB that mediate addiction (see the opening story of this chapter) are involved with histone acetylation, which leads to subsequent gene activation. The sperm of men with psychosis have different methylation patterns than sperm from nonpsychotic men. This last observation is especially provocative, as it suggests that epigenetic patterns, some of which may have formed during life, can be passed on to the next generation. This means that some phenotypic characteristics acquired during the lifetime of an organism might be heritable, contrary to biologists’ long-held views. The idea that epigenetic changes can be inherited remains controversial.

DNA methylation can result in genomic imprinting

In mammals specific patterns of methylation develop for each sex during gamete formation. This happens in two stages: first, the existing methyl groups are removed from the 5-methylcytosines by a demethylase, and then a DNA methylase adds methyl groups to the appropriate cytosines. When the gametes form they carry this new pattern of methylation (epigenetic information).

The DNA methylation pattern in male gametes (sperm) differs from that in female gametes (eggs) at about 200 genes in the mammalian genome. That is, a given gene in this group may be methylated in eggs but unmethylated in sperm (**Figure 16.20**). In this case the offspring would inherit a maternal gene that is transcriptionally inactive (methylated) and a paternal gene that is transcriptionally active (demethylated). This is called **genomic imprinting**.



16.20 Genomic Imprinting For some genes, epigenetic DNA methylation differs in male and female gametes. As a result, an individual might inherit an allele from the female parent that is transcriptionally silenced; but the same allele from the male parent would be expressed.

An example of imprinting is found in a region on human chromosome 15 called 15q11. This region is imprinted differently during the formation of male and female gametes, and offspring normally inherit both the paternally and maternally derived patterns. In rare cases, there is a chromosome deletion in one of the gametes, and the newborn baby inherits just the male or the female imprinting pattern in this particular chromosome region. If the male pattern is the only one present (female region deleted), the baby develops Angelman Syndrome, characterized by epilepsy, tremors, and constant smiling. If the female pattern is the only one present (male region deleted), the baby develops a quite different phenotype called Prader-Willi syndrome, marked by muscle weakness and obesity. Note that the *gene sequences are the same* in both cases: it is *the epigenetic patterns that are different*.

Imprinting of specific genes occurs primarily in mammals and flowering plants. Most imprinted genes are involved with embryonic development. An embryo must have both the paternally and maternally imprinted gene patterns to develop properly. In fact, attempts to make an embryo that has chromosomes from only one sex (for example, by chemically treating an egg cell to double its chromosomes) usually fail. So imprinting has an important lesson for genetics: *males and females may be the same genetically (except for the X and Y chromosomes), but they differ epigenetically.*

Global chromosome changes involve DNA methylation

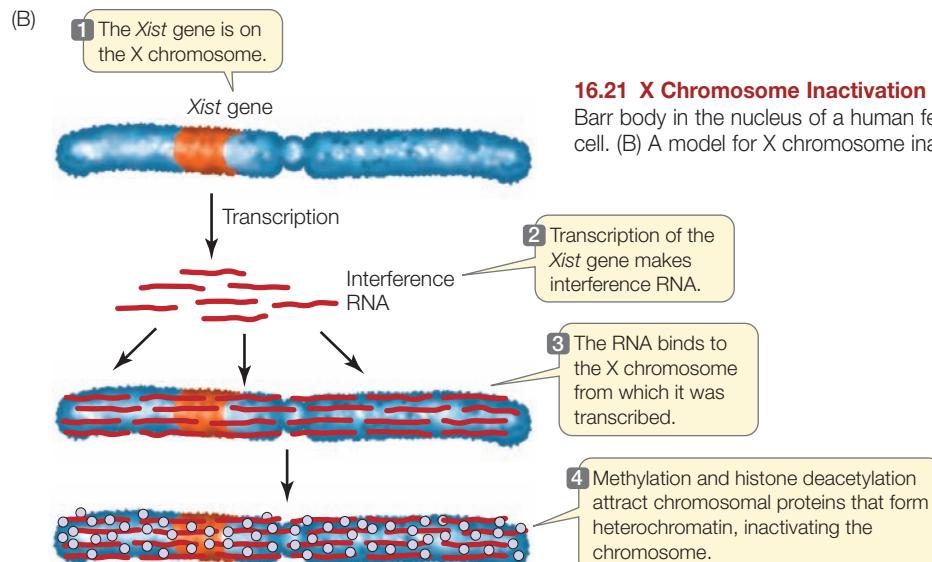
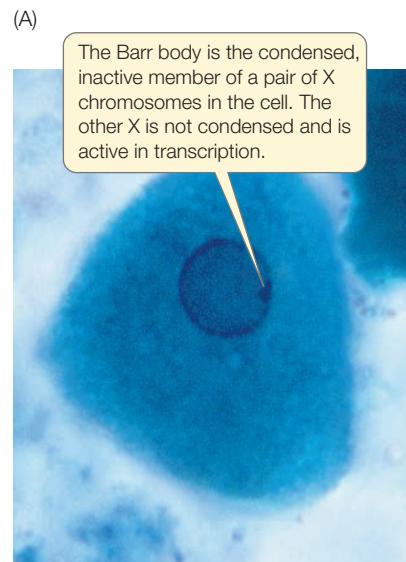
Like single genes, large regions of chromosomes or even entire chromosomes can have distinct patterns of DNA methylation. Under a microscope, two kinds of chromatin can be distinguished in the stained interphase nucleus: *euchromatin* and *heterochromatin*. The euchromatin appears diffuse and stains lightly; it contains the DNA that is transcribed into mRNA. Heterochromatin is condensed and stains darkly; any genes it contains are generally not transcribed.

Perhaps the most dramatic example of heterochromatin is the inactive X chromosome of mammals. A normal female mammal has two X chromosomes; a normal male has an X and a Y (see Section 12.4). The X and Y chromosomes probably arose from a pair of autosomes (non-sex chromosomes) about 300 million years ago. Over time, mutations in the Y chromosome resulted in maleness-determining genes, and the Y chromosome gradually lost most of the genes it once shared with its X homolog. As a result, females and males differ greatly in the “dosage” of X-linked genes. Each female cell has two copies of each gene on the X chromosome, and therefore has the potential to produce twice as much of each protein product. Nevertheless, for 75 percent of the genes on the X chromosome, transcription is generally the same in males and in females. How does this happen?

Mary Lyon, Liane Russell, and Ernest Beutler independently hypothesized in 1961 that one of the X chromosomes in each cell of a female is, to a significant extent, transcriptionally inactivated early in embryonic development. They proposed that one copy of X becomes inactive in each embryonic cell, and the same X remains inactive in all that cell’s descendants. Several lines of evidence have since confirmed this hypothesis.

In a given embryonic cell, the “choice” of which X in the pair to inactivate is random. Recall that one X in a female comes from her father and one from her mother. Thus, in one embryonic cell the paternal X might be the one remaining transcriptionally active, but in a neighboring cell the maternal X might be active.

The inactivated X chromosome does not vanish, but is identifiable within the nucleus. During interphase a single, stainable nuclear body called a Barr body (after its discoverer, Murray Barr) can be seen in cells of human females under the light microscope (Figure 16.21A). This clump of heterochromatin, which is not present in normal males, is the inactivated X chromosome, and it consists of heavily methylated DNA. A female with the normal two X chromosomes will have one Barr body, while a rare female with three Xs will have two, and an XXXX female will have three. Males that are XYY will have one. These observations suggest that the interphase cells of each person, male



16.21 X Chromosome Inactivation (A) A Barr body in the nucleus of a human female cell. (B) A model for X chromosome inactivation.

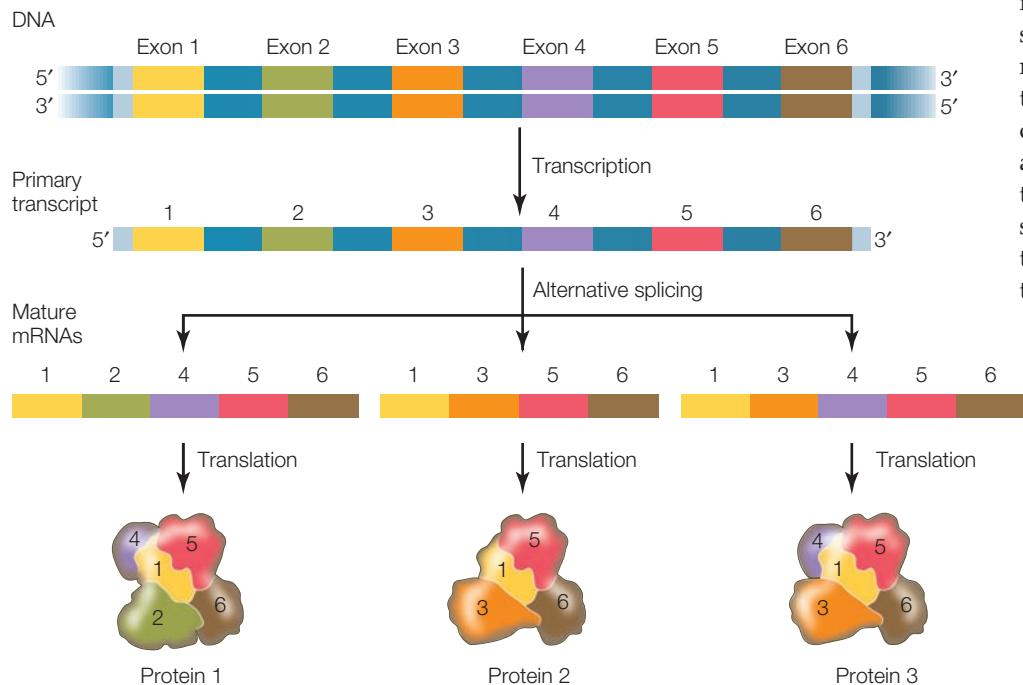
or female, have a single active X chromosome, and thus a constant dosage of expressed X chromosome genes.

Condensation of the inactive X chromosome makes its DNA sequences physically unavailable to the transcriptional machinery. Most of the genes of the inactive X are heavily methylated. However, one gene, *Xist* (for X inactivation-specific transcript), is only lightly methylated and is transcriptionally active. On the active X chromosome, *Xist* is heavily methylated and not transcribed. The RNA transcribed from *Xist* binds to the X chromosome from which it is transcribed, and this binding leads to a spreading of inactivation along the chromosome. The *Xist* RNA transcript is an example of **interference RNA** (Figure 16.21B).

16.4 RECAP

Epigenetics describes stable changes in gene expression that do not involve changes in DNA sequences. These changes involve modifications of DNA (cytosine methylation) or of histone proteins bound to DNA. Epigenetic changes can be affected by the environment, and can also result in genome imprinting, in which expression of some genes depends on their parental origin.

- How are DNA methylation patterns established and how do they affect gene expression? See p. 356 and Figure 16.18
- Explain how histone modifications affect transcription. See pp. 357–358 and Figure 16.19
- Why and how does X chromosome inactivation occur? See p. 359



Gene expression involves transcription and then translation. So far we have described how eukaryotic gene expression is regulated at the transcriptional level. But as Figure 16.13 shows, there are many points at which regulation can occur after the initial gene transcript is made.

16.5 How Is Eukaryotic Gene Expression Regulated After Transcription?

Eukaryotic gene expression can be regulated both in the nucleus prior to mRNA export, and after the mRNA leaves the nucleus. Posttranslational control mechanisms can involve alternative splicing of pre-mRNA, microRNAs, repressors of translation, or regulation of protein breakdown in the proteasome.

Different mRNAs can be made from the same gene by alternative splicing

Most primary mRNA transcripts contain several introns (see Figure 14.7). We have seen how the splicing mechanism recognizes the boundaries between exons and introns. What would happen if the β -globin pre-mRNA, which has two introns, were spliced from the start of the first intron to the end of the second? The middle exon would be spliced out along with the two introns. An entirely new protein (certainly not a β -globin) would be made, and the functions of normal β -globin would be lost. Such **alternative splicing** can be a deliberate mechanism for generating a family of different proteins with different activities and functions from a single gene (Figure 16.22).

Before the human genome was sequenced, most scientists estimated that they would find between 80,000 and 150,000 protein-coding genes. You can imagine their surprise when the actual sequence revealed only about 24,000 genes! In fact, there are many more human mRNAs than there are human genes, and most of this variation comes from alternative splicing. Indeed, recent surveys show that about half of all human genes are alternatively spliced. Alternative splicing may be a key to the differences in levels of complexity among organisms. For example, although humans and chimpanzees have similar-sized genomes, there is more alternative splicing in the human brain than in the brain of a chimpanzee.

genes, and most of this variation comes from alternative splicing. Indeed, recent surveys show that about half of all human genes are alternatively spliced. Alternative splicing may be a key to the differences in levels of complexity among organisms. For example, although humans and chimpanzees have similar-sized genomes, there is more alternative splicing in the human brain than in the brain of a chimpanzee.

16.22 Alternative Splicing Results in Different Mature mRNAs and Proteins

Pre-mRNA can be spliced differently in different tissues, resulting in different proteins.

MicroRNAs are important regulators of gene expression

As we discuss in the next chapter, less than 5 percent of the genome in most plants and animals codes for proteins. Some of the genome encodes ribosomal RNA and transfer RNAs, but until recently biologists thought that the rest of the genome was not transcribed; some even called it “junk.” Recent investigations, however, have shown that some of these noncoding regions are transcribed. The noncoding RNAs are often very small and therefore difficult to detect. These tiny RNA molecules are called **microRNA** (miRNA).

The first miRNA sequences were found in the worm *Caenorhabditis elegans*. This model organism, which has been studied extensively by developmental biologists, goes through several larval stages. Victor Ambros at the University of Massachusetts found mutations in two genes that had different effects on progress through these stages:

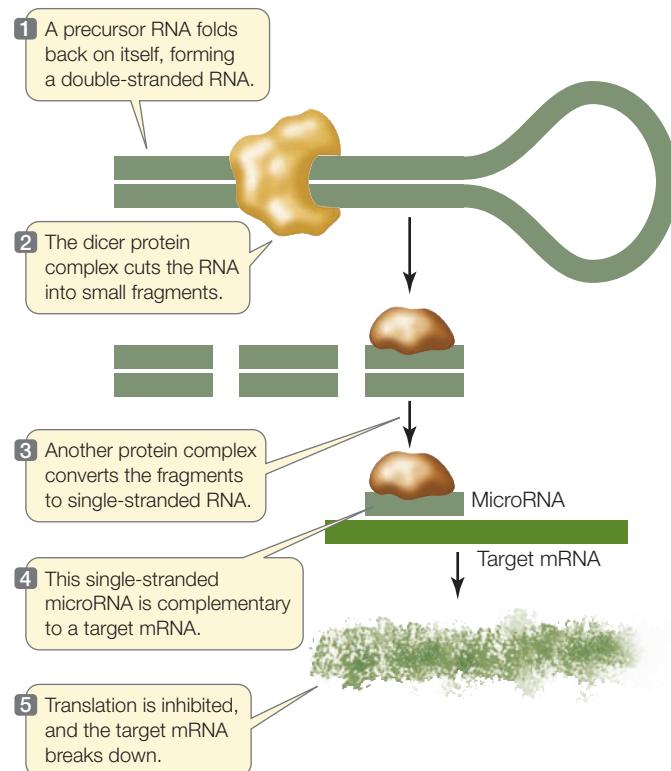
- *lin-14* mutations (named for abnormal cell *lineage*) caused the larvae to skip the first stage and go straight to the second stage. Thus the gene’s normal role is to facilitate events of the first larval stage.
- *lin-4* mutations caused certain cells in later larval stages to repeat a pattern of development normally shown in the first larval stage. It was as if the cells were stuck in that stage. So the normal role of this gene is to *negatively regulate lin-14*, turning its expression off so the cells can progress to the next stage.

Not surprisingly, further investigation showed that *lin-14* encodes a transcription factor that affects the transcription of genes involved in larval cell progression. It was originally expected that *lin-4*, the negative regulator, would encode a protein that downregulates genes activated by the *lin-14* protein. But this turned out to be incorrect. Instead, *lin-4* encodes a 22-base miRNA that inhibits *lin-14* expression posttranscriptionally by binding to its mRNA.

Several hundred miRNAs have now been described in many eukaryotes. Each one is about 22 bases long and usually has dozens of mRNA targets. These miRNAs are transcribed as longer precursors that are then cleaved through a series of steps to double-stranded miRNAs. A protein complex guides the miRNA to its target mRNA, where translation is inhibited and the mRNA is degraded (Figure 16.23). The remarkable conservation of the miRNA gene silencing mechanism in eukaryotes indicates that it is evolutionarily ancient and biologically important.

Translation of mRNA can be regulated

Is the amount of a protein in a cell determined by the amount of its mRNA? Recently, scientists examined the relationship between mRNA abundance and protein abundance in yeast cells. For about a third of the many genes surveyed, there was a clear correlation between mRNA and protein: more of one led to more of the other. But for two-thirds of the proteins, there was no apparent relationship between the two: sometimes there was lots of mRNA and little or no protein, or lots of protein and lit-



16.23 mRNA Inhibition by MicroRNAs MicroRNAs result in inhibition of translation and in breakdown of the target mRNA.

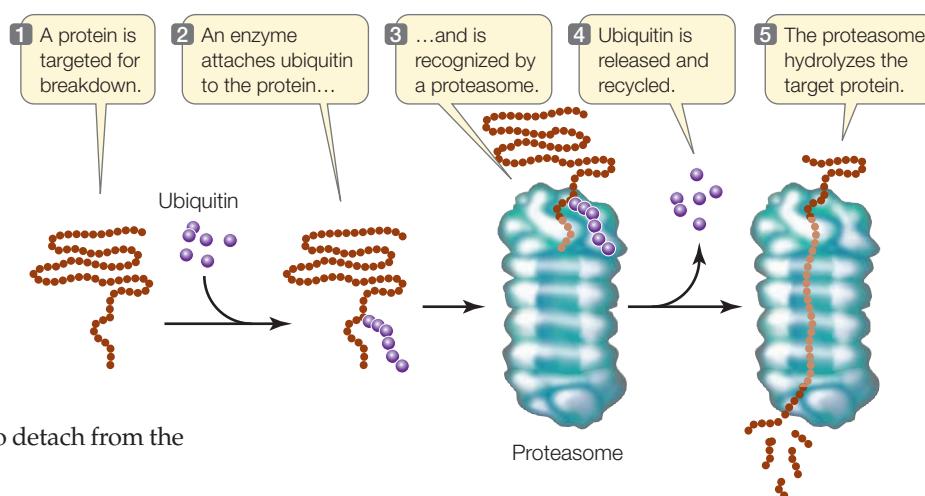
tle mRNA. The concentrations of these proteins must therefore be determined by factors acting after the mRNA is made. Cells do this in two major ways: by blocking the translation of mRNA, or by altering how long newly synthesized proteins persist in the cell (protein longevity).

REGULATION OF TRANSLATION There are three known ways in which the translation of mRNA can be regulated. One way, as we saw in the previous section, is to inhibit translation with miRNAs. A second way involves modification of the guanosine triphosphate cap on the 5' end of the mRNA (see Section 14.4). An mRNA that is capped with an unmodified GTP molecule is not translated. For example, stored mRNAs in the egg cells of the tobacco hornworm moth are capped with unmodified GTP molecules and are not translated. After the egg is fertilized, however, the caps are modified, allowing the mRNA to be translated to produce the proteins needed for early embryonic development.

In another system, repressor proteins directly block translation. For example, in mammalian cells the protein ferritin binds free iron ions (Fe^{2+}). When iron is present in excess, ferritin synthesis rises dramatically, but the amount of ferritin mRNA remains constant, indicating that the increase in ferritin synthesis is due to an increased rate of mRNA translation. Indeed, when the iron level in the cell is low, a translational repressor protein binds to ferritin mRNA and prevents its translation by blocking its attachment to a ribosome. When the iron level rises, some of the excess Fe^{2+} ions bind to the repressor and alter its three-

16.24 A Proteasome Breaks Down Proteins

Proteins targeted for degradation are bound to ubiquitin, which then binds the targeted protein to a proteasome. The proteasome is a complex structure where proteins are digested by several powerful proteases.



dimensional structure, causing the repressor to detach from the mRNA and allowing translation to proceed.

REGULATION OF PROTEIN LONGEVITY The protein content of any cell at a given time is a function of both protein synthesis and protein degradation. Certain proteins can be targeted for destruction in a chain of events that begins when an enzyme attaches a 76-amino acid protein called **ubiquitin** (so named because it is ubiquitous, or widespread) to a lysine residue of the protein to be destroyed. Other ubiquitins then attach to the primary one, forming a polyubiquitin chain. The protein–polyubiquitin complex then binds to a huge protein complex called a **proteasome** (from *protease* and *soma*, body) (Figure 16.24). Upon entering the proteasome, the polyubiquitin is removed and ATP energy is used to unfold the target protein. Three different proteases then digest the protein into small peptides and amino acids.

You may recall from Section 11.2 that cyclins are proteins that regulate the activities of key enzymes at specific points in the cell cycle. Cyclins must be broken down at just the right time, and this is done by proteasomes. Viruses can hijack this system. For example, some strains of the human papillomavirus target p53 protein and retinoblastoma protein, which normally inhibit the cell cycle, for proteasomal degradation, resulting in unregulated cell division (cancer).

16.5 RECAP

One of the most important means of posttranscriptional regulation is alternative RNA splicing, which allows more than one protein to be made from a single gene. The stability of mRNA in the cytoplasm can also be regulated. MicroRNAs, mRNA modifications, and translational repressors can prevent mRNA translation. Proteins in the cell can be targeted for breakdown by ubiquitin and then hydrolyzed in proteasomes.

- How can a single pre-mRNA sequence encode several different proteins? See p. 360 and Figure 16.22
- How do miRNAs regulate gene expression? See p. 361 and Figure 16.23
- Explain the role of the proteasome. See p. 362 and Figure 16.24

CHAPTER SUMMARY

16.1 How Do Viruses Regulate Their Gene Expression?

- **Viruses** are not cells, and rely on host cells to reproduce.
- The basic unit of a virus is a **virion**, which consists of a nucleic acid genome (DNA or RNA) and a protein coat, called a **capsid**.
- **Bacteriophage** are viruses that infect bacteria.
- Viruses undergo a **lytic cycle**, which causes the host cell to burst, releasing new virions.
- Some viruses have promoters that bind host RNA polymerase, which they use to transcribe their own genes and proteins. **Review Figure 16.3**
- Rarely, a phage will transfer bacterial genes to a new host in the process of **transduction**.
- Some viruses can also undergo a **lysogenic cycle**, in which a molecule of their DNA, called a **prophage**, is inserted into the host chromosome, where it replicates for generations. **Review Figure 16.4**

- The cellular environment determines whether a phage undergoes a lytic or a lysogenic cycle. Regulatory proteins that compete for promoters on phage DNA control the switch between the two life cycles. **Review Figure 16.5**

- A **retrovirus** uses reverse transcriptase to generate a cDNA **provirus** from its RNA genome. The provirus is incorporated into the host's DNA and can be activated to produce new virions. **Review Figure 16.6**

16.2 How Is Gene Expression Regulated in Prokaryotes?

- Some proteins are synthesized only when they are needed. Proteins that are made only in the presence of a particular compound—an **inducer**—are **inducible** proteins. Proteins that are made at a constant rate regardless of conditions are **constitutive** proteins.
- An **operon** consists of a promoter, an **operator**, and two or more **structural genes**. Promoters and operators do not code

for proteins, but serve as binding sites for regulatory proteins. [Review Figure 16.10](#)

- **Regulatory genes** code for regulatory proteins, such as **repressors**. When a repressor binds to an operator, transcription of the structural gene is inhibited. [Review Figure 16.11](#), [ANIMATED TUTORIALS 16.1 AND 16.2](#)
- The *lac* operon is an example of an inducible system, in which the presence of an inducer (lactose) keeps the repressor from binding the operator, allowing the transcription of structural genes for lactose metabolism.
- Transcription can be enhanced by the binding of an **activator** protein to the promoter. [Review Figure 16.12](#)
- **Catabolite repression** is the inhibition of a catabolic pathway for one energy source by a different, preferred energy source.

16.3 How Is Eukaryotic Gene Transcription Regulated?

- Eukaryotic gene expression is regulated both during and after transcription. [Review Figure 16.13](#), [WEB ACTIVITY 16.1](#)
- **Transcription factors** and other proteins bind to DNA and affect the rate of initiation of transcription at the promoter. [Review Figure 16.14 and 16.15](#), [ANIMATED TUTORIAL 16.3](#)
- The interactions of these proteins with DNA are highly specific and depend on protein domains and DNA sequences.
- Genes at distant locations from one another can be coordinately regulated by common transcription factors and promoter elements. [Review Figure 16.17](#)

16.4 How Do Epigenetic Changes Regulate Gene Expression?

- **Epigenetics** refers to changes in gene expression that do not involve changes in DNA sequences.
- **Methylation** of cytosine residues generally inhibits transcription. [Review Figure 16.18](#)
- Modifications of histone proteins in nucleosomes make transcription either easier or more difficult. [Review Figure 16.19](#)
- Epigenetic changes can occur because of the environment.
- DNA methylation can explain **genome imprinting**, where the expression of a gene depends on its parental origin. [Review Figure 16.20](#)

16.5 How Is Eukaryotic Gene Expression Regulated After Transcription?

- **Alternative splicing** of pre-mRNA can produce different proteins. [Review Figure 16.22](#)
- **MicroRNAs** are small RNAs that do not code for proteins, but regulate the translation and longevity of mRNA. [Review Figure 16.23](#)
- The translation of mRNA to proteins can be regulated by translational repressors.
- The **proteasome** can break down proteins, thus affecting protein longevity. [Review Figure 16.24](#)

SEE WEB ACTIVITY 16.2 for a concept review of this chapter.

SELF-QUIZ

1. Which of the following statements about the *lac* operon is *not* true?
 - a. When lactose binds to the repressor, the repressor can no longer bind to the operator.
 - b. When lactose binds to the operator, transcription is stimulated.
 - c. When the repressor binds to the operator, transcription is inhibited.
 - d. When lactose binds to the repressor, the shape of the repressor is changed.
 - e. The repressor has binding sites for both DNA and lactose.
2. Which of the following is *not* a type of viral reproduction?
 - a. DNA virus in a lytic cycle
 - b. DNA virus in a lysogenic cycle
 - c. DNA virus (single-stranded) with a double-stranded DNA intermediate
 - d. RNA virus with reverse transcription to make cDNA
 - e. RNA virus acting as tRNA
3. In the lysogenic cycle of bacteriophage λ ,
 - a. a repressor, cl, blocks the lytic cycle.
 - b. the bacteriophage carries DNA between bacterial cells.
 - c. both early and late phage genes are transcribed.
 - d. the viral genome is made into RNA, which stays in the host cell.
 - e. many new viruses are made immediately, regardless of host health.
4. An operon is
 - a molecule that can turn genes on and off.
 - b. an inducer bound to a repressor.
 - c. a series of regulatory sequences controlling transcription of protein-coding genes.
5. Which of the following is true of both positive and negative gene regulation?
 - a. They reduce the rate of transcription of certain genes.
 - b. They involve regulatory proteins (or RNA) binding to DNA.
 - c. They involve transcription of all genes in the genome.
 - d. They are not both active in the same organism or virus.
 - e. They act away from the promoter.
6. In DNA, 5-methylcytosine
 - a. forms a base pair with adenine.
 - b. is not recognized by DNA polymerase.
 - c. is related to transcriptional silencing of genes.
 - d. does not occur at promoters.
 - e. is an irreversible modification of cytosine.
7. Which statement about selective gene transcription in eukaryotes is *not* true?
 - a. Regulatory proteins can bind at a site on DNA distant from the promoter.
 - b. Transcription requires transcription factors.
 - c. Genes are usually transcribed as groups called operons.
 - d. Both positive and negative regulation occur.
 - e. Many proteins bind at the promoter.
8. Control of gene expression in eukaryotes includes all of the following except
 - a. alternative RNA splicing.
 - b. binding of proteins to DNA.
 - c. transcription factors.
 - d. stabilization of mRNA by miRNA.
 - e. DNA methylation.

9. The promoter in the *lac* operon is
 - a. the region that binds the repressor.
 - b. the region that binds RNA polymerase.
 - c. the gene that codes for the repressor.
 - d. a structural gene.
 - e. an operon.
10. Epigenetic changes
 - a. can involve DNA methylation.
 - b. are due to nonhistone protein acetylation.
 - c. are due to changes in the genetic code.
 - d. are an example of positive control of translation.
 - e. are never reversible.

FOR DISCUSSION

1. Compare the life cycles of a lysogenic bacteriophage and HIV (Figures 16.4 and 16.6) with respect to:
 - a. how the virus enters the cell.
 - b. how the virion is released in the cell.
 - c. how the viral genome is replicated.
 - d. how new viruses are produced.
2. Compare promoters adjacent to early and late genes in the bacteriophage lytic cycle.
3. The repressor protein that acts on the *lac* operon of *E. coli* is encoded by a regulatory gene. The repressor is made in small quantities and at a constant rate. Would you surmise that the promoter for this repressor protein is efficient or inefficient? Is synthesis of the repressor constitutive, or is it under environmental control?
4. A protein-coding gene in a eukaryote has three introns. How many different proteins could be made by alternative splicing of the pre-mRNA from this gene?

ADDITIONAL INVESTIGATION

In colorectal cancer, tumor suppressor genes are not active. This is an important factor resulting in uncontrolled cell division. Two possible explanations for the inactive genes are: a mutation in the coding region, resulting in an inactive protein,

or epigenetic silencing at the promoter of the gene, resulting in reduced transcription. How would you investigate these two possibilities?

17 Genomes

The dog genome

Canis lupus familiaris, the dog, was domesticated by humans from the gray wolf thousands of years ago. While there are many kinds of wolves, they all look more or less the same. Not so with “man’s best friend.” The American Kennel Club recognizes about 155 different breeds. Dog breeds not only look different, they vary greatly in size. For example, an adult Chihuahua weighs just 1.5 kg, while a Scottish deerhound weighs 70 kg. No other mammal shows such large phenotypic variation, and biologists are curious about how this occurs. Also, there are hundreds of genetic diseases in dogs, and many of these dis-

eases have counterparts in humans. To find out about the genes behind the phenotypic variation, and to elucidate the relationships between genes and diseases, the Dog Genome Project began in the late 1990s. Since then the sequences of several dog genomes have been published.

Two dogs—a boxer and a poodle—were the first to have their entire genomes sequenced. The dog genome contains 2.8 billion base pairs of DNA in 39 pairs of chromosomes. There are 19,000 protein-coding genes, most of them with close counterparts in other mammals, including humans. The whole genome sequence made it easy to create a map of genetic markers—specific nucleotides or short sequences of DNA at particular locations on the genome that differ between individual dogs and/or breeds.

Genetic markers are used to map the locations of (and thus identify) genes that control particular traits. For example, Dr. Elaine Ostrander and her colleagues at the National Institutes of Health studied Portuguese water dogs to identify genes that control size. Taking samples of cells for DNA isolation was relatively easy: a cotton swab was swept over the inside of the cheek. As Dr. Ostrander said, the dogs “didn’t care, especially if they thought they were going to get a treat or if there was a tennis ball in our other hand.” It turned out that the gene for insulin-like growth factor 1 (IGF-1) is important in determining size: large breeds have an allele that codes for an active IGF-1 and small breeds have a different allele that codes for a less active IGF-1.

Another gene important to phenotypic variation was found in whippets, sleek dogs that run fast and are often raced. A mutation in the gene for myostatin, a protein that inhibits overdevelopment of muscles, results in a



Variation in Dogs The Chihuahua (bottom) and the Brazilian mastiff (top) are the same species, *Canis lupus familiaris*, and yet show great variation in size. Genome sequencing has revealed insights into how size is controlled by genes.



Genetic Bully These dogs are both whippets, but the muscle-bound dog (right) has a mutation in a gene that limits muscle buildup.

whippet that is more muscular and runs faster. Myostatin is important in human muscles as well.

Inevitably, some scientists have set up companies to test dogs for genetic variations, using DNA supplied by anxious owners and breeders. Some traditional breeders frown on this practice, but others say it will improve the breeds and give more joy (and prestige) to owners. So the issues surrounding the Dog Genome Project are not very different from ones arising from the Human Genome Project.

Powerful methods have been developed to analyze DNA sequences, and the resulting information is accumulating at a rapid rate. Comparisons of sequenced genomes are providing new insights into evolutionary relationships and confirming old ones. We are in a new era of biology.

IN THIS CHAPTER we look at genomes. First we look at how large molecules of DNA are cut and sequenced, and what kinds of information these genome sequences provide. Then we turn to the results of ongoing sequencing efforts in both prokaryotes and eukaryotes. We next consider the human genome and some of the real and potential uses of human genome information. Finally, we will describe the emerging fields of proteomics and metabolomics, which attempt to give a complete inventory of a cell's proteins and metabolic activity.

CHAPTER OUTLINE

- 17.1 How Are Genomes Sequenced?**
- 17.2 What Have We Learned from Sequencing Prokaryotic Genomes?**
- 17.3 What Have We Learned from Sequencing Eukaryotic Genomes?**
- 17.4 What Are the Characteristics of the Human Genome?**
- 17.5 What Do the New Disciplines of Proteomics and Metabolomics Reveal?**

17.1 How Are Genomes Sequenced?

As you saw in the opening story on dogs, one reason for sequencing genomes is to compare different organisms. Another is to identify changes in the genome that result in disease. In 1986, the Nobel laureate Renato Dulbecco and others proposed that the world scientific community be mobilized to undertake the sequencing of the entire human genome. One challenge discussed at the time was to detect DNA damage in people who had survived the atomic bomb attacks and been exposed to radiation in Japan during World War II. But in order to detect changes in the human genome, scientists first needed to know its normal sequence.

The result was the publicly funded **Human Genome Project**, an enormous undertaking that was successfully completed in 2003. This effort was aided and complemented by privately funded groups. The project benefited from the development of many new methods that were first used in the sequencing of smaller genomes—those of prokaryotes and simple eukaryotes.

Two approaches were used to sequence the human genome

Many prokaryotes have a single chromosome, while eukaryotes have several to many. Because of their differing sizes, chromosomes can be separated from one another, identified, and experimentally manipulated. It might seem that the most straightforward approach to sequencing a chromosome would be to start at one end and simply sequence the entire DNA molecule. However, this approach is not practical since only about 700 base pairs can be sequenced at a time using current methods. Prokaryotic chromosomes contain 1–4 million base pairs and human chromosome 1 contains 246 million base pairs.

To sequence an entire genome, chromosomal DNA must be cut into short fragments about 500 base pairs long, which are separated and sequenced. For the haploid human genome, which has about 3.3 billion base pairs, there are more than 6 million such fragments. When all of the fragments have been sequenced, the problem becomes how to put these millions of sequences together. This task can be accomplished using larger, overlapping fragments.

Let's illustrate this process using a single, 10 base-pair (bp) DNA molecule. (This is a double-stranded molecule, but for convenience we show only the sequence of the noncoding

strand.) The molecule is cut three ways. The first cut generates the fragments:

TG, ATG, and CCTAC

The second cut of the same molecule generates the fragments:

AT, GCC, and TACTG

The third cut results in:

CTG, CTA, and ATGC

Can you put the fragments into the correct order? (The answer is ATGCCTACTG.) Of course, the problem of ordering 6 million fragments, each about 500 bp long, is more of a challenge! The field of **bioinformatics** was developed to analyze DNA sequences using complex mathematics and computer programs.

Until recently, two broad approaches were used to analyze DNA fragments for alignment: hierarchical sequencing and shotgun sequencing. These were developed for the Human Genome Project, but have been applied to other organisms as well.

HIERARCHICAL SEQUENCING The publicly funded human genome sequencing team developed a method known as **hierarchical sequencing**. The first step was to systematically identify short marker sequences along the chromosomes, ensuring that every fragment of DNA to be sequenced would contain a marker (**Figure 17.1A**). Genetic markers can be short tandem repeats (STRs), single nucleotide polymorphisms (SNPs), or the recognition sites for *restriction enzymes*, which recognize and cut DNA at specific sequences (see Chapter 15).

Some restriction enzymes recognize sequences of 4–6 base pairs and generate many fragments from a large DNA molecule. For example, the enzyme *Sau3A* cuts DNA every time it encounters GATC. Other restriction enzymes recognize sequences of 8–12 base pairs (*NotI* cuts at GCGGCCGC, for example) and generate far fewer, but much larger, fragments.

In hierarchical sequencing, genomic DNA is cut up into a set of relatively large (55,000 to 2 million bp) fragments. If different enzymes are used in separate digests, the fragments will overlap so that some fragments share particular markers. Each fragment is inserted into a bacterial plasmid to create a **bacterial artificial chromosome (BAC)**, which is then inserted into bacteria. Each bacterium gets just one plasmid with its fragment of (for example) the human genome and is allowed to grow into a colony containing millions of genetically identical bacteria (called a *clone*). Clones differ from one another in that each has a different fragment from the human genome. A collection of clones, containing many different fragments of a genome, is called a **genomic library**.

The DNA from each clone is then extracted and cut into smaller overlapping pieces, which in turn are cloned, purified, and sequenced. The overlapping parts of the sequences allow researchers (with the aid of computers) to align them to create the complete sequence of the BAC clone. The genetic markers on each BAC clone are used to arrange the larger fragments in the proper order along the chromosome map. This method works, but it is slow. An alternative approach, shotgun sequencing, makes far greater use of use of computers to align the sequences.

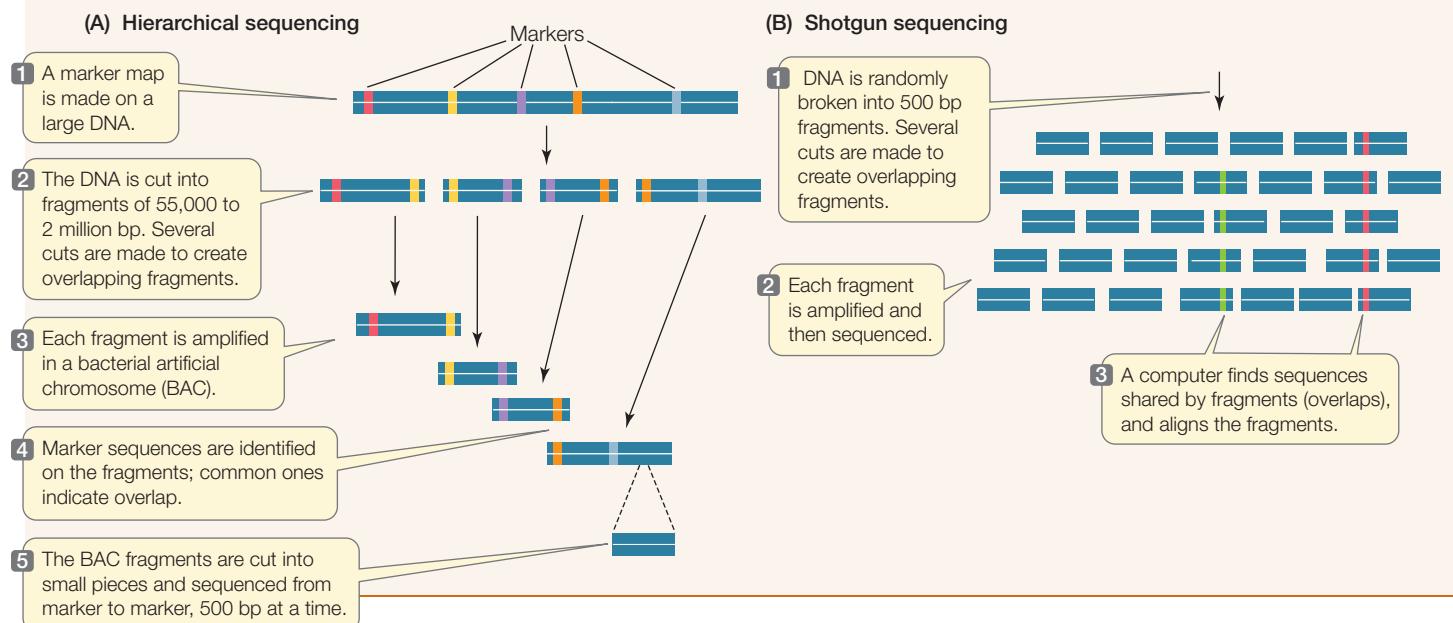
yourBioPortal.com

GO TO Animated Tutorial 17.1 • Sequencing the Genome

TOOLS FOR INVESTIGATING LIFE

17.1 Sequencing Genomes Involves Fragment Overlaps

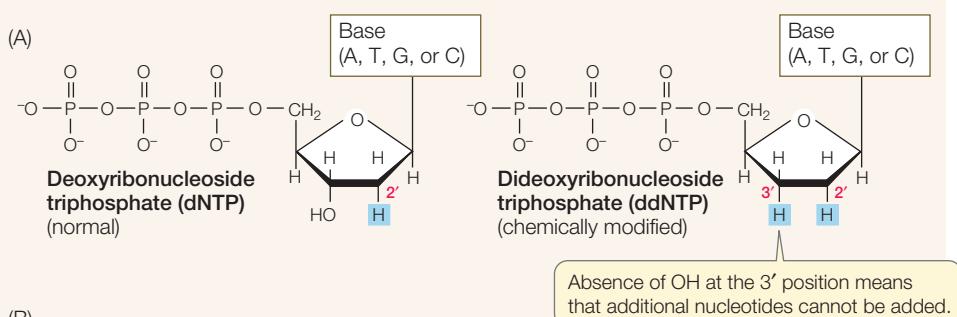
Short fragments of the whole genome can be sequenced, but then the fragments must be correctly aligned. Historically two approaches were used. Both involved the use of bacterial clones to separate and amplify individual DNA fragments.



TOOLS FOR INVESTIGATING LIFE

17.2 Sequencing DNA

(A) The normal substrates for DNA replication are dNTPs. The chemically modified structure of ddNTPs causes DNA synthesis to stop. (B) When labeled ddNTPs are incorporated into a reaction mixture for replicating a DNA template of unknown sequence, the result is a collection of fragments of varying lengths that can be separated by electrophoresis.



- 1** The DNA fragment for which the base sequence is to be determined is isolated and serves as the template.

2 Each of the ddNTPs is bound to a fluorescent dye.

3 A sample of the unknown DNA is combined with primer, DNA polymerase, dNTPs, and the fluorescent ddNTPs. Synthesis begins.

4 The results are illustrated here by what binds to a T in the unknown template. If ddATP A is added, synthesis stops. A series of fragments of different lengths is made, each ending with a ddNTP.

5 The newly synthesized fragments of various lengths are separated by electrophoresis.

6 Each fragment fluoresces a color that identifies the ddNTP that terminated the fragment. The color at the end of each fragment is detected by a laser beam.

7 The sequence of the DNA can now be deduced from the colors of each fragment...

8 ...and converted to the sequence of the template strand.

Template strand
5' ??????????????????CGCA 3'
3' GCGT 5' (Primer sequence known)

5' T?????????????????CGCA 3'
3' ATCTGGCTATTGGGGCGT 5'

5' TT?????????????????CGCA 3'
3' ATCTGGCTATTGGGGCGT 5'

Electrophoresis 3'
A Longest fragment
T
C
T
G
G
G
C
T
A
T
C
G Shortest fragment
G
G
5'

Laser
Detector

3' A A T C T G G G C T A T T C G G 5'

5' T T A G A C C C G A T A G C C C G C A 3'

SHOTGUN SEQUENCING Instead of mapping the genome and creating a BAC library, the **shotgun sequencing** method involves directly cutting genomic DNA into smaller, overlapping fragments that are cloned and sequenced. Powerful computers align the fragments by finding sequence homologies in the overlapping regions (**Figure 17.1B**). As sequencing technologies and computers have improved, the shotgun approach has become much faster and cheaper than the hierarchical approach.

As a demonstration, researchers used this method to sequence a 1.8 million-base-pair prokaryotic genome in just a few months. Next came larger genomes. The entire 180 million-base-pair fruit fly genome was sequenced by the shotgun method in little over a year. This success proved that the shotgun method might work for the much larger human genome, and in fact it was used to sequence the human genome rapidly relative to the hierarchical method.

The nucleotide sequence of DNA can be determined

How are the individual DNA fragments generated by the hierarchical or shotgun methods sequenced? Current techniques are variations of a method developed in the late 1970s by Frederick Sanger. This method uses chemically modified nucleotides that were originally developed to stop cell division in cancer. As we discuss in Chapter 13, deoxyribonucleoside triphosphates (dNTPs) are the normal substrates for DNA replication, and contain the sugar deoxyribose. If that sugar is replaced with 2,3-dideoxyribose, the resulting dideoxyribonucleoside triphosphate (ddNTP) will still be added by DNA polymerase to a growing polynucleotide chain. However, because the ddNTP has no hydroxyl group ($-\text{OH}$) at the 3' position, the next nucleotide cannot be added (Figure 17.2A). Thus synthesis stops at the position where ddNTP has been incorporated into the growing end of a DNA strand.

To determine the sequence of a DNA fragment (usually no more than 700 base pairs long), it is isolated and mixed with

- DNA polymerase
- A short primer appropriate for the DNA sequence
- The four dNTPs (dATP, dGTP, dCTP, and dTTP)
- Small amounts of the four ddNTPs, each bonded to a differently colored fluorescent “tag”

In the first step of the reaction, the DNA is heated to denature it (separate it into single strands). Only one of these strands will act as a template for sequencing—the one to which the primer binds. DNA replication proceeds, and the test tube soon contains a mixture of the original DNA strands and shorter, new complementary strands. The new strands, each ending with a fluorescent ddNTP, are of varying lengths. For example, each time a T is reached on the template strand, DNA polymerase adds either a dATP or a ddATP to the growing complementary strand. If dATP is added, the strand continues to grow. If ddATP is added, growth stops (Figure 17.2B).

After DNA replication has been allowed to proceed for a while, the new DNA fragments are denatured and the single-stranded fragments separated by electrophoresis (see Figure 15.8), which sorts the DNA fragments by length. During the electrophoresis run, the fragments pass through a laser beam that excites the fluorescent tags, and the distinctive color of light emitted by each ddNTP is detected. The color indicates which ddNTP is at the end of each strand. A computer processes this information and prints out the DNA sequence of the fragment (see Figure 17.2B).

The delivery of chemical reagents by automated machines, coupled with automated analysis, has made DNA sequencing faster than ever. Huge laboratories often have 80 sequencing machines operating at once, each of which can sequence and analyze up to 70,000 bp in a typical 4-hour run. This may be fast enough for a prokaryotic genome with 1.5 million base pairs (20 runs), but when it comes to routine sequencing of larger genomes (like the 3.3 billion-base-pair human genome), even more speed is needed.

High-throughput sequencing has been developed for large genomes

The first decade of the new millennium has seen rapid development of **high-throughput sequencing** methods—fast, cheap ways to sequence and analyze large genomes. A variety of different approaches are being used. They generally involve the amplification of DNA templates by the polymerase chain reaction (PCR; see Section 13.5), and the physical binding of template DNA to a solid surface or to tiny beads called microbeads. These

techniques are often referred to as *massively parallel DNA sequencing*, because thousands or millions of sequencing reactions are run at once to greatly speed up the process. One such high-throughput method is illustrated in Figure 17.3. In one 7-hour run, these machines can sequence 50,000,000 base pairs of DNA! How does it work?

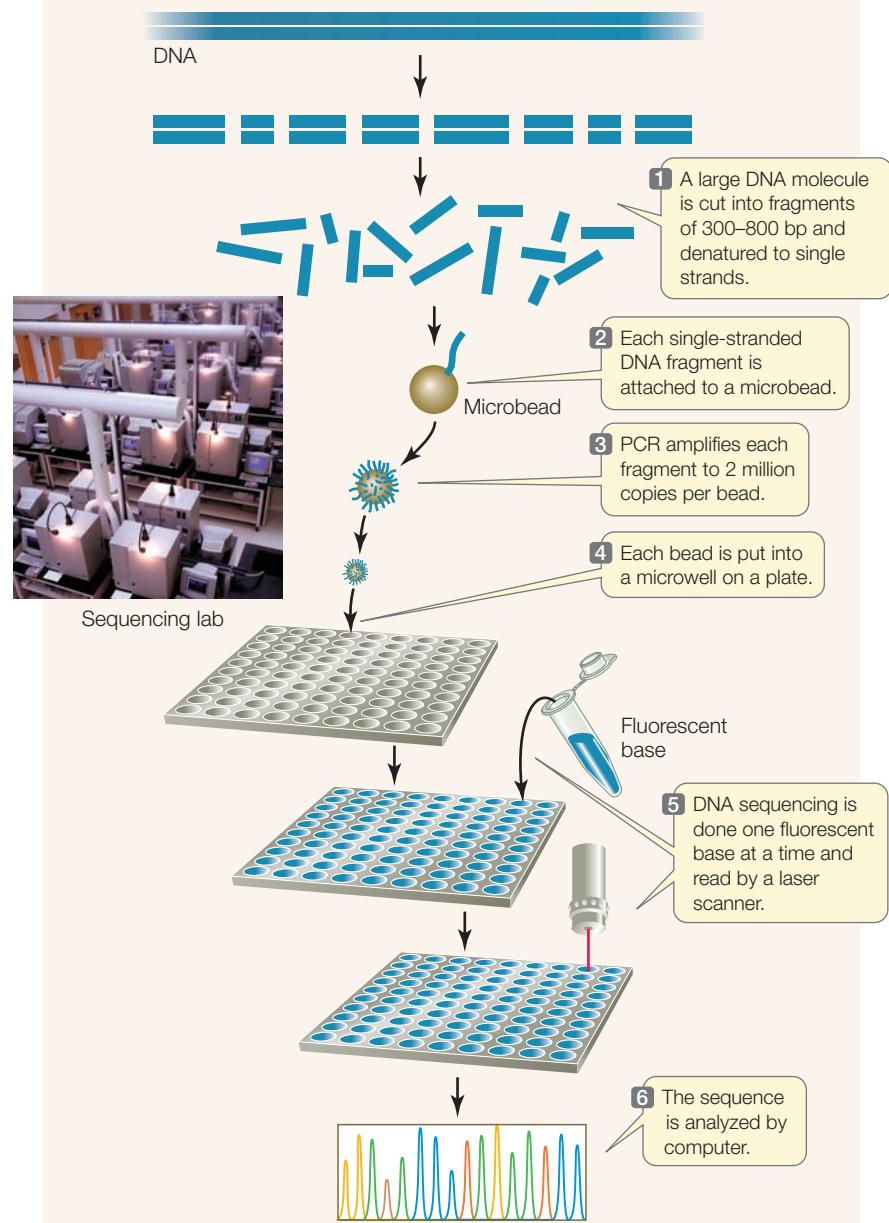
yourBioPortal.com

GO TO Animated Tutorial 17.2 • High-Throughput Sequencing

TOOLS FOR INVESTIGATING LIFE

17.3 High-Throughput Sequencing

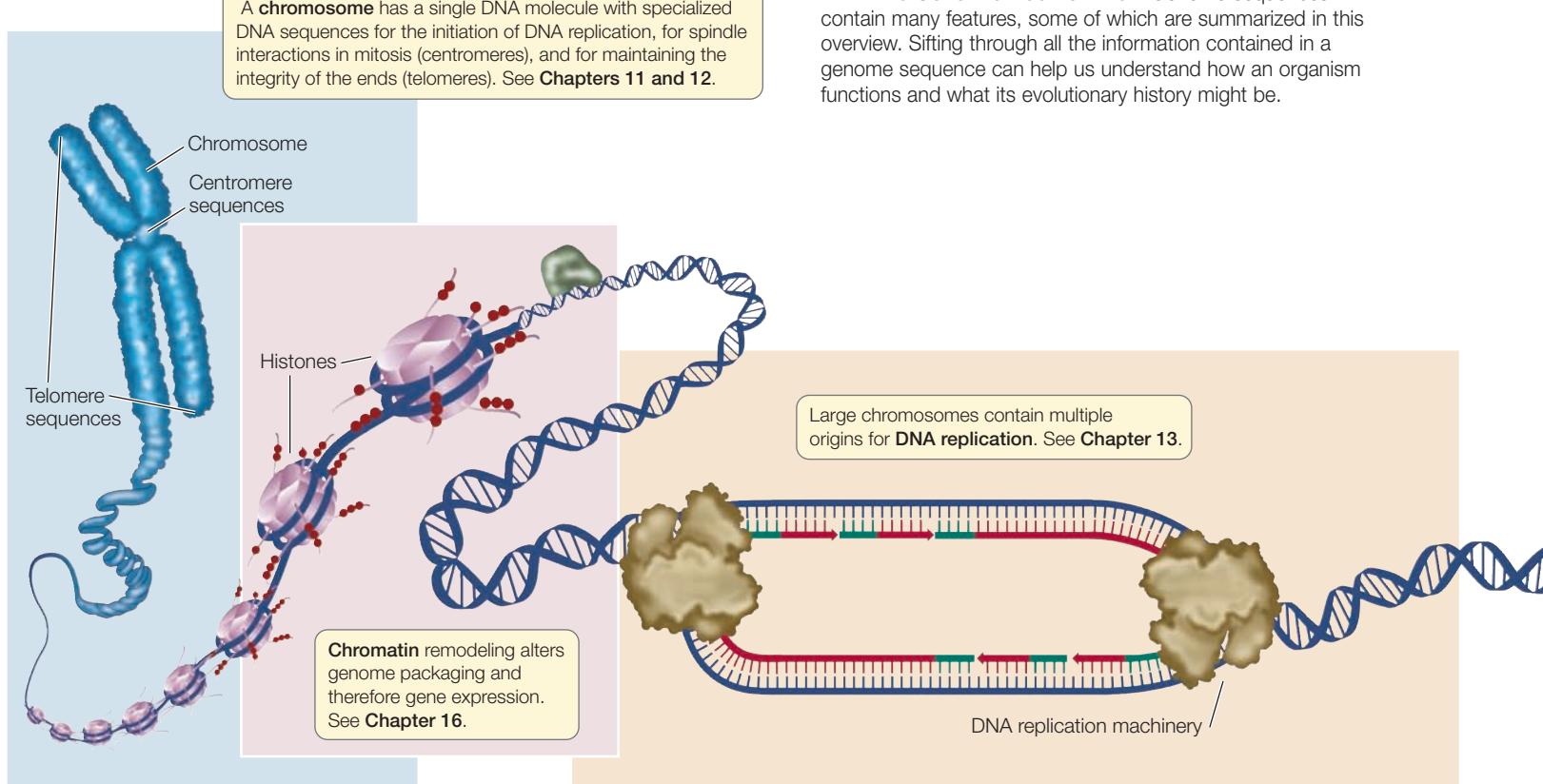
High-speed sequencing is faster and cheaper than traditional methods, and involves the chemical amplification of DNA fragments. One example of high-throughput sequencing is shown here.



For massively parallel sequencing using microbeads, the genomic DNA is first cut into 300- to 800-base-pair fragments. The fragments are denatured to single strands and attached to tiny beads that are less than 20 μm in diameter, one DNA fragment (template) per bead. PCR is used to create several million identical copies of the fragment on each bead. Then each bead is loaded into a tiny (40 μm diameter) well in a multi-well plate, and the sequencing begins.

The automated sequencer adds a reaction mix like the one described above, but containing only one of four fluorescently labeled dNTPs. That nucleotide will become incorporated as the first nucleotide in a complementary strand only in wells where the first nucleotide in the template strand can base-pair with it. For example, if the first nucleotide on the template in well #1 has base T, then a fluorescent nucleotide with base A will bind to that well. Next, the reaction mix is removed and a scanner captures an image of the plate, indicating which wells contain the fluorescent nucleotide. This process is repeated with a different labeled nucleotide. The machine cycles through many repeats using all four dNTPs, and records which wells gain new nucleotides after each cycle. A computer then identifies the sequence of nucleotides that were gained by each well, and aligns the fragments to provide the complete sequence of the genome.

This method was used to sequence the genome of James Watson, codiscoverer of the DNA double helix. It took less than two months and cost less than \$1 million. Sequencing methods are being continually refined to increase speed and accuracy and decrease costs.



Genome sequences yield several kinds of information

New genome sequences are published more and more frequently, creating a torrent of biological information (Figure 17.4). In general, biologists use sequence information to identify:

- *Open reading frames*, the coding regions of genes. For protein-coding genes, these regions can be recognized by the start and stop codons for translation, and by intron consensus sequences that indicate the locations of introns.
- *Amino acid sequences* of proteins, which can be deduced from the DNA sequences of open reading frames by applying the genetic code (see Figure 14.6).
- *Regulatory sequences*, such as promoters and terminators for transcription.
- *RNA genes*, including rRNA, tRNA, and small nuclear RNA (snRNA) genes.
- *Other noncoding sequences* that can be classified into various categories including centromeric and telomeric regions, nuclear matrix attachment regions, transposons, and repetitive sequences such as short tandem repeats.

Sequence information is also used for *comparative genomics*, the comparison of a newly sequenced genome (or parts thereof) with sequences from other organisms. This can give information about the functions of sequences, and can be used to trace evolutionary relationships among different organisms.

17.1 RECAP

The sequencing of genomes required the development of ways to cut large chromosomes into fragments, sequence the fragments, and then line them up on the chromosome. Two ways to do this are hierarchical sequencing and shotgun sequencing. Today new procedures are being developed that require automation and powerful computers. Actual DNA sequencing involves labeled nucleotides that are detected at the ends of growing polynucleotide chains.

- What are the hierarchical and shotgun approaches to genome analysis? **See pp. 367–368 and Figure 17.1**
- What is the dideoxy method for DNA sequencing? **See pp. 368–369 and Figure 17.2**
- Explain how high-throughput sequencing methods work. **See pp. 369–370 and Figure 17.3**
- How are open reading frames recognized in a genomic sequence? What kind of information can be derived from an open reading frame? **See p. 370**

We now turn to the first organisms whose sequences were determined, prokaryotes, and the information these sequences provided.

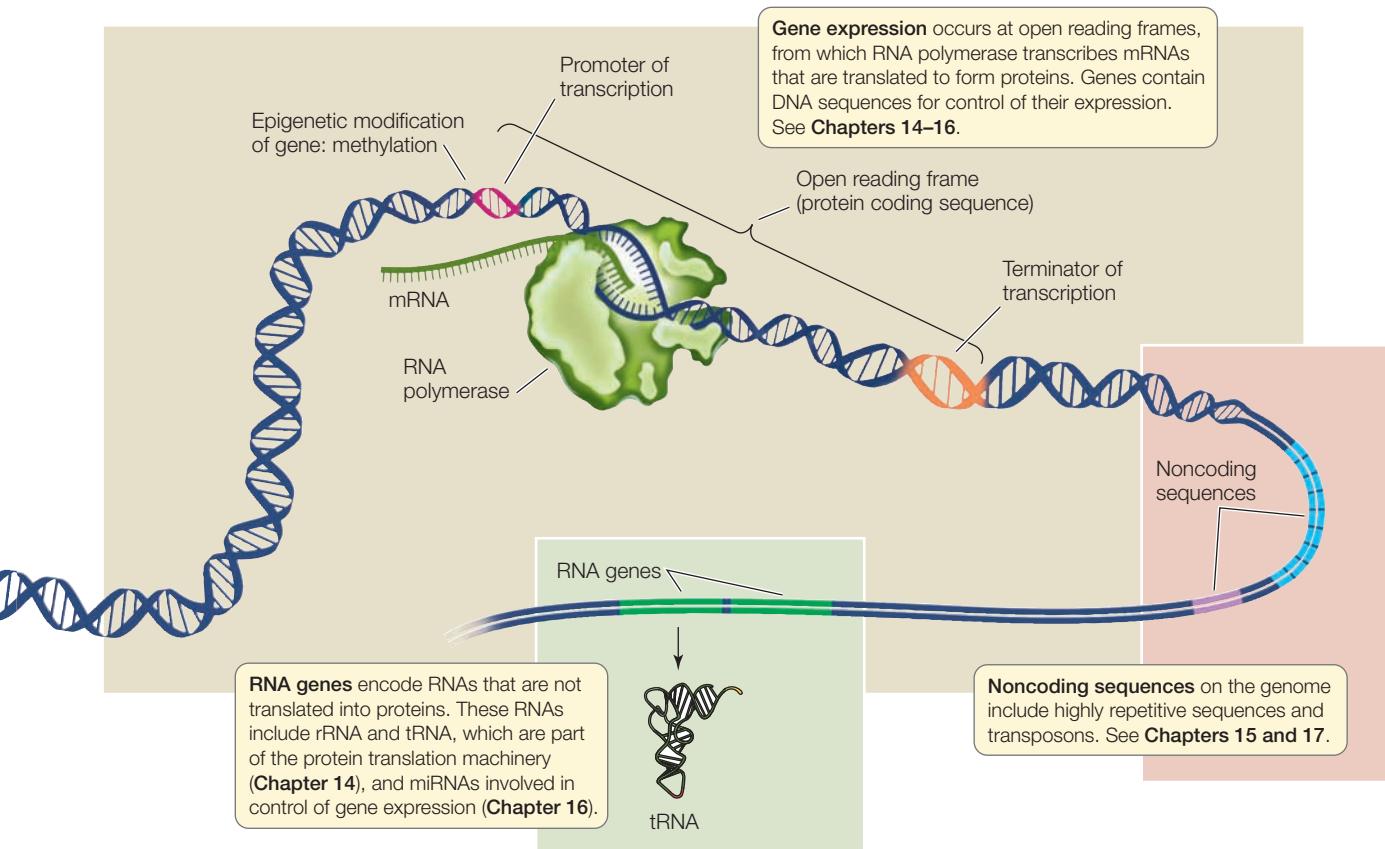
17.2 What Have We Learned from Sequencing Prokaryotic Genomes?

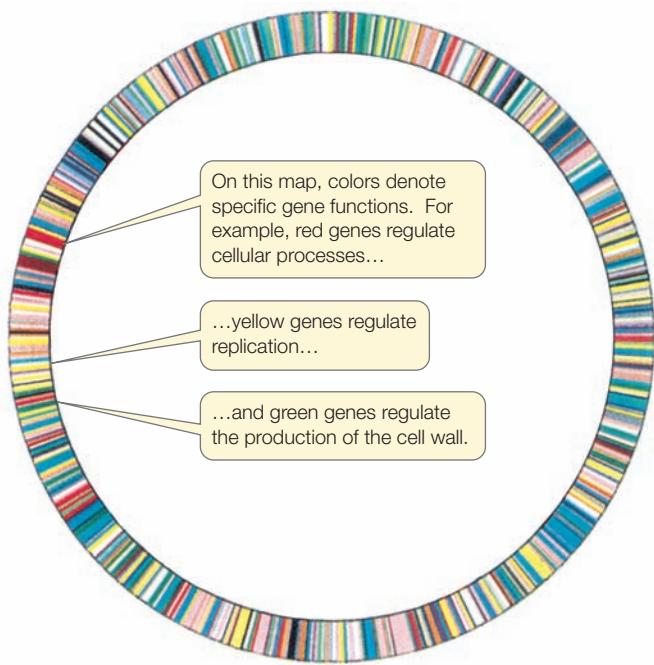
When DNA sequencing became possible in the late 1970s, the first life forms to be sequenced were the simplest viruses with their relatively small genomes. The sequences quickly provided new information on how these viruses infect their hosts and reproduce. But the manual sequencing techniques used on viruses were not up to the task of studying the genomes of prokaryotes and eukaryotes. The newer, automated sequencing techniques we just described made such studies possible. We now have genome sequences for many prokaryotes, to the great benefit of microbiology and medicine.

The sequencing of prokaryotic genomes led to new genomics disciplines

In 1995 a team led by Craig Venter and Hamilton Smith determined the first complete genomic sequence of a free-living cellular organism, the bacterium *Haemophilus influenzae*. Many more prokaryotic sequences have followed, revealing not only how prokaryotes apportion their genes to perform different cellular functions, but also how their specialized functions are carried out. Soon we may even be able to ask the provocative question of what the minimal requirements of a living cell might be.

FUNCTIONAL GENOMICS Functional genomics is the biological discipline that assigns functions to the products of genes. This





17.5 Functional Organization of the Genome of *H. influenzae*
The entire DNA sequence has 1,830,137 base pairs. Different colors reflect different classes of gene function.

field, less than 15 years old, is now a major occupation of biologists. Let's see how functional genomics methods were applied to the bacterium *H. influenzae* once its sequence was known.

The only host for *H. influenzae* is humans. It lives in the upper respiratory tract and can cause ear infections or, more seriously, meningitis in children. Its single circular chromosome has 1,830,138 base pairs (Figure 17.5). In addition to its origin of replication and the genes coding for rRNAs and tRNAs, this bacterial chromosome has 1,738 open reading frames with promoters nearby.

When this sequence was first announced, only 1,007 (58 percent) of the open reading frames coded for proteins with known functions. The remaining 42 percent coded for proteins whose functions were unknown. Since then scientists have identified many of these proteins' roles. For example, they found genes for enzymes of glycolysis, fermentation, and electron transport. Other gene sequences code for membrane proteins, including those involved in active transport. An important finding was that highly infective strains of *H. influenzae*, but not noninfective strains, have genes for surface proteins that attach the bacterium to the human respiratory tract. These surface proteins are now a focus of research on possible treatments for *H. influenzae* infections.

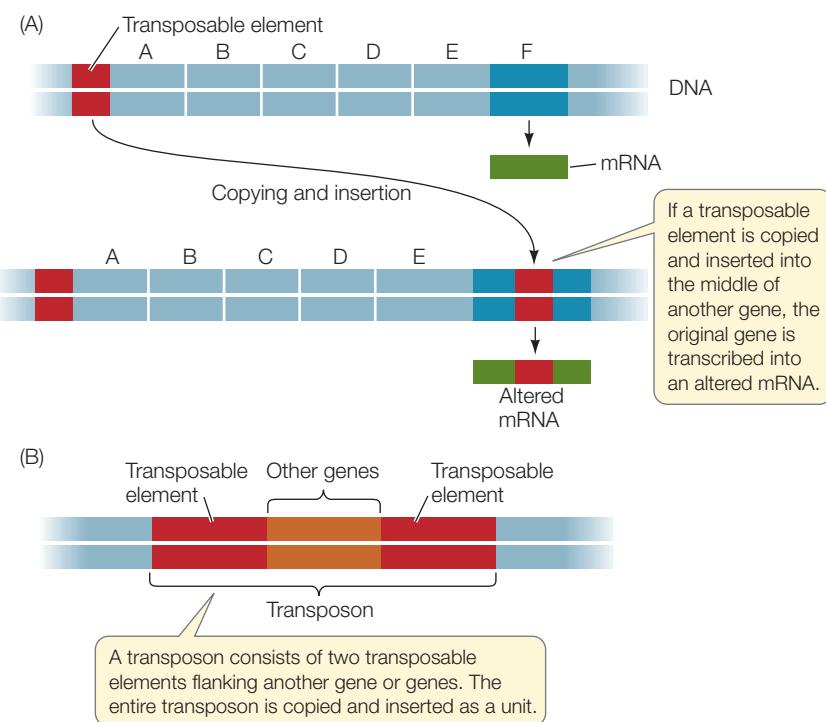
COMPARATIVE GENOMICS Soon after the sequence of *H. influenzae* was announced, smaller (*Mycoplasma genitalium*; 580,073 base pairs) and larger (*E. coli*; 4,639,221 base pairs) prokaryotic sequences were completed. Thus began a new era in biology, that of **comparative genomics**,

which compares genome sequences from different organisms. Scientists can identify genes that are present in one bacterium and missing in another, allowing them to relate these genes to bacterial function.

M. genitalium, for example, lacks the enzymes needed to synthesize amino acids, which *E. coli* and *H. influenzae* both possess. This finding reveals that *M. genitalium* must obtain all its amino acids from its environment (usually the human urogenital tract). Furthermore, *E. coli* has 55 regulatory genes coding for transcriptional activators and 58 for repressors; *M. genitalium* only has 3 genes for activators. What do such findings tell us about an organism's lifestyle? For example, is the biochemical flexibility of *M. genitalium* limited by its relative lack of control over gene expression?

Some sequences of DNA can move about the genome

Genome sequencing allowed scientists to study more broadly a class of DNA sequences that had been discovered by geneticists decades earlier. Segments of DNA called **transposable elements** can move from place to place in the genome and can even be inserted into another piece of DNA in the same cell (e.g., a plasmid). A transposable element might be at one location in the genome of one *E. coli* strain, and at a different location in another strain. The insertion of this movable DNA sequence from elsewhere in the genome into the middle of a protein-coding gene disrupts that gene (Figure 17.6A). Any mRNA expressed from the disrupted gene will have the extra sequence and the



17.6 DNA Sequences that Move Transposable elements are DNA sequences that move from one location to another. (A) In one method of transposition, the DNA sequence is replicated and the copy inserts elsewhere in the genome. (B) Transposons contain transposable elements and other genes.

protein will be abnormal. So transposable elements can produce significant phenotypic effects by inactivating genes.

Transposable elements are often short sequences of 1,000–2,000 base pairs, and are found at many sites in prokaryotic genomes. The mechanisms that allow them to move vary. For example, a transposable element may be replicated, and then the copy inserted into another site in the genome. Or the element might splice out of one location and move to another location.

Longer transposable elements (up to 5,000 bp) carry additional genes and are called **transposons** (Figure 17.6B). Sometimes these DNA regions contain a gene for antibiotic resistance.

The sequencing of prokaryotic and viral genomes has many potential benefits

Prokaryotic genome sequencing promises to provide insights into microorganisms that cause human diseases. Genome sequencing has revealed unknown genes and proteins that can be targeted for isolation and functional study. Such studies are revealing new methods to combat pathogens and their infections. Sequencing has also revealed surprising relationships between some pathogenic organisms, suggesting that genes may be transferred between different strains.

- *Chlamydia trachomatis* causes the most common sexually transmitted disease in the United States. Because it is an intracellular parasite, it has been very hard to study. Among its 900 genes are several for ATP synthesis—something scientists used to think this bacterium could not accomplish on its own.
- *Rickettsia prowazekii* causes typhus; it is carried by lice and infects people bitten by the lice. Of its 634 genes, 6 encode proteins that are essential for virulence. These virulence proteins are being used to develop vaccines.
- *Mycobacterium tuberculosis* causes tuberculosis. It has a relatively large genome, coding for 4,000 proteins. Over 250 of these are used to metabolize lipids, so this may be the main way that this bacterium gets its energy. Some of its genes code for previously unidentified cell surface proteins; these proteins are targets for potential vaccines.
- *Streptomyces coelicolor* and its close relatives are the source for the genes for two-thirds of all naturally occurring antibiotics currently in clinical use. These antibiotics include streptomycin, tetracycline, and erythromycin. The genome sequence of *S. coelicolor* reveals 22 clusters of genes responsible for antibiotic production, of which only four were previously known. This finding may lead to new antibiotics to combat pathogens that have evolved resistance to conventional antibiotics.
- *E. coli* strain O157:H7 causes illness (sometimes severe) in at least 70,000 people a year in the United States. Its genome has 5,416 genes, of which 1,387 are different from those in the familiar (and harmless) laboratory strains of this bacterium. Many of these unique genes are also present in other pathogenic bacteria, such as *Salmonella* and *Shigella*. This finding suggests that there is extensive genetic ex-

change among these species, and that “superbugs” that share genes for antibiotic resistance may be on the horizon.

- *Severe acute respiratory syndrome (SARS)* was first detected in southern China in 2002 and rapidly spread in 2003. There is no effective treatment and 10 percent of infected people die. Isolation of the causative agent, a virus, and the rapid sequencing of its genome revealed several novel proteins that are possible targets for antiviral drugs or vaccines. Research is underway on both fronts, since another outbreak is anticipated.

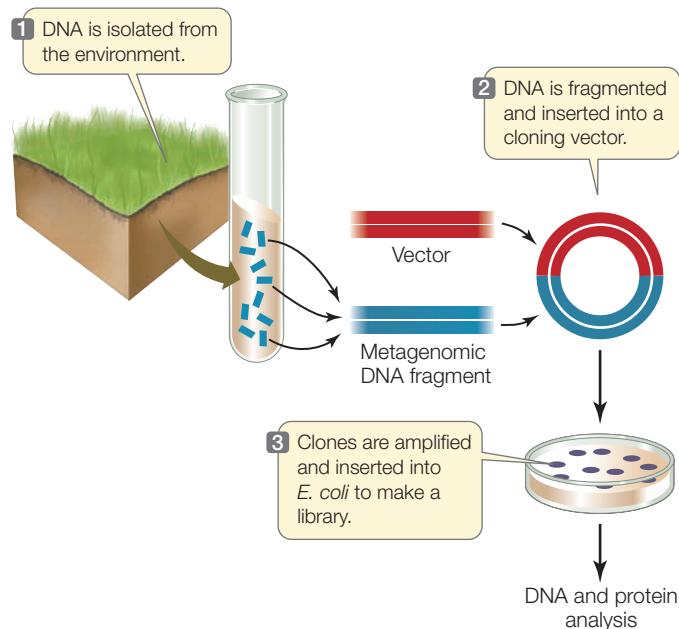
Genome sequencing also provides insights into organisms involved in global ecological cycles (see Chapter 58). In addition to the well-known carbon dioxide, another important gas contributing to the atmospheric “greenhouse effect” and global warming is methane (CH_4 ; see Figure 2.7). Some bacteria, such as *Methanococcus*, produce methane in the stomachs of cows. Others, such as *Methylococcus*, remove methane from the air and use it as an energy source. The genomes of both of these bacteria have been sequenced. Understanding the genes involved in methane production and oxidation may help us to slow the progress of global warming.

Metagenomics allows us to describe new organisms and ecosystems

If you take a microbiology laboratory course you will learn how to identify various prokaryotes on the basis of their growth in lab cultures. For example, staphylococci are a group of bacteria that infect skin and nasal passages. When grown on a special medium called blood agar they form round, raised colonies. Microorganisms can also be identified by their nutritional requirements or the conditions under which they will grow (for example, aerobic versus anaerobic). Such culture methods have been the mainstay of microbial identification for over a century and are still useful and important. However, scientists can now use PCR and modern DNA analysis techniques to analyze microbes *without* culturing them in the laboratory.

In 1985, Norman Pace, then at Indiana University, came up with the idea of isolating DNA directly from environmental samples. He used PCR to amplify specific sequences from the samples to determine whether particular microbes were present. The PCR products were sequenced to explore their diversity. The term **metagenomics** was coined to describe this approach of analyzing genes without isolating the intact organism. It is now possible to perform shotgun sequencing with samples from almost any environment. The sequences can be used to detect the presence of known microbes and pathogens, and perhaps even the presence of heretofore unidentified organisms (Figure 17.7). For example:

- Shotgun sequencing of DNA from 200 liters of seawater indicated that it contained 5,000 different viruses and 2,000 different bacteria, many of which had not been described previously.
- One kilogram of marine sediment contained a million different viruses, most of them new.



17.7 Metagenomics Microbial DNA extracted from the environment can be amplified and analyzed. This has led to the description of many new genes and species.

- Water runoff from a mine contained many new species of prokaryotes thriving in this apparently inhospitable environment. Some of these organisms exhibited metabolic pathways that were previously unknown to biologists. These organisms and their capabilities may be useful in cleaning up pollutants from the water.

These and other discoveries are truly extraordinary and potentially very important. It is estimated that 90 percent of the microbial world has been invisible to biologists and is only now being revealed by metagenomics. Entirely new ecosystems of bacteria and viruses are being discovered in which, for example, one species produces a molecule that another metabolizes. It is hard to overemphasize the importance of such an increase in our knowledge of the hidden world of microbes. This knowledge will help us to understand natural ecological processes, and has the potential to help us find better ways to manage environmental catastrophes such as oil spills, or remove toxic heavy metals from soil.

Will defining the genes required for cellular life lead to artificial life?

When the genomes of prokaryotes and eukaryotes are compared, a striking conclusion arises: certain genes are present in all organisms (universal genes). There are also some (nearly) universal gene segments that are present in many genes in many organisms; for example, the sequence that codes for an ATP binding site. These findings suggest that there is some ancient, minimal set of DNA sequences common to all cells. One way to identify these sequences is to look for them in computer analyses of sequenced genomes.

Another way to define the minimal genome is to take an organism with a simple genome and deliberately mutate one gene at a time to see what happens. *M. genitalium* has one of the smallest known genomes—only 482 protein-coding genes. Even so, some of its genes are dispensable under some circumstances. For example, it has genes for metabolizing both glucose and fructose, but it can survive in the laboratory on a medium containing only one of these sugars.

What about other genes? Researchers have addressed this question with experiments involving the use of transposons as mutagens. When transposons in the bacterium are activated, they insert themselves into genes at random, mutating and inactivating them (Figure 17.8). The mutated bacteria are tested for growth and survival, and DNA from interesting mutants is sequenced to find out which genes contain transposons.

The astonishing result of these studies is that *M. genitalium* can survive in the laboratory with a minimal genome of only 382 functional genes! Is this really all it takes to make a viable organism? Experiments are underway to make a synthetic genome based on that of *M. genitalium*, and then insert it into an empty bacterial cell. If the cell starts transcribing mRNA and making proteins—is in fact viable—it may turn out to be the first life created by humans.

In addition to the technical feat of creating artificial life, this technique could have important applications. New microbes could be made with entirely new abilities, such as degrading oil spills, making synthetic fibers, reducing tooth decay, or converting cellulose to ethanol for use as fuel. On the other hand, fears of the misuse or mishandling of this knowledge are not unfounded. For example, it might also be possible to develop synthetic bacteria harmful to people, animals or plants, and use them as agents of biological warfare or bioterrorism. The “genomics genie” is, for better or worse, already out of the bottle. Hopefully human societies will use it to their benefit.

17.2 RECAP

DNA sequencing is used to study the genomes of prokaryotes that are important to humans and to ecosystems. Functional genomics uses gene sequences to determine the functions of the gene products. Comparative genomics compares gene sequences from different organisms to help identify their functions and evolutionary relationships. Transposable elements and transposons move from one place to another in the genome.

- Give some examples of prokaryotic genomes that have been sequenced. What have the sequences shown? See pp. 371–373
- What is metagenomics and how is it used? See pp. 373–374 and Figure 17.7
- How are selective inactivation studies being used to determine the minimal genome? See p. 374 and Figure 17.8

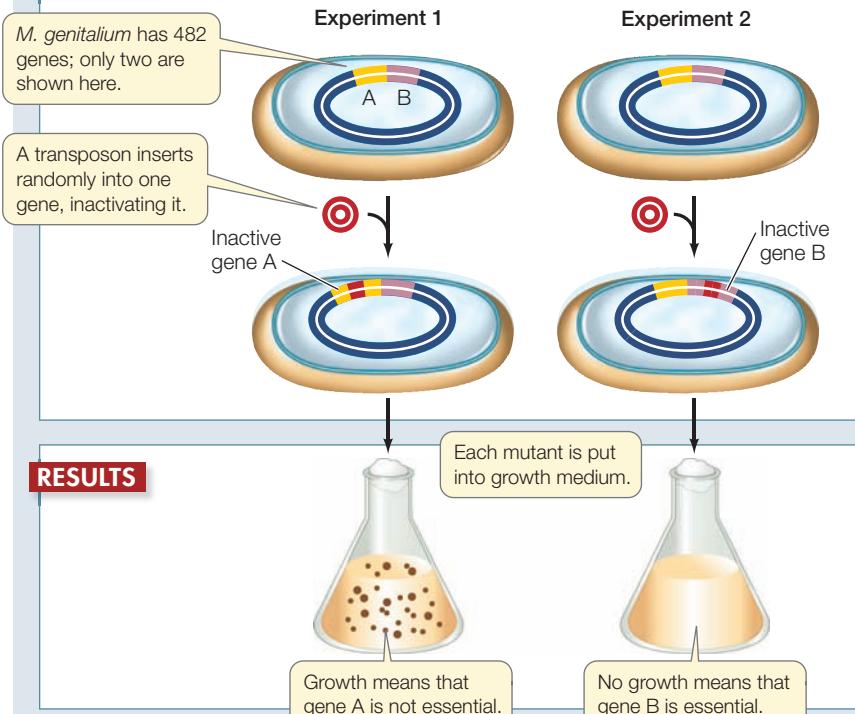
INVESTIGATING LIFE

17.8 Using Transposon Mutagenesis to Determine the Minimal Genome

Mycoplasma genitalium has the smallest number of genes of any prokaryote. But are all of its genes essential to life? By inactivating the genes one by one, scientists determined which of them are essential for the cell's survival. This research may lead to the construction of artificial cells with customized genomes, designed to perform functions such as degrading oil and making plastics.

HYPOTHESIS Only some of the genes in a bacterial genome are essential for cell survival.

METHOD



CONCLUSION If each gene is inactivated in turn, a "minimal essential genome" can be determined.

Go to yourBioPortal.com for original citations, discussions, and relevant links for all INVESTIGATING LIFE figures.

Advances in DNA sequencing and analysis have led to the rapid sequencing of eukaryotic genomes. We now turn to the results of these analyses.

17.3 What Have We Learned from Sequencing Eukaryotic Genomes?

As genomes have been sequenced and described, a number of major differences have emerged between eukaryotic and prokaryotic genomes (Table 17.1). Key differences include:

- Eukaryotic genomes are larger than those of prokaryotes, and they have more protein-coding genes. This difference is not surprising, given that multicellular organisms have many cell types with specific functions. Many proteins are needed

to do those specialized jobs. A typical virus contains enough DNA to code for only a few proteins—about 10,000 base pairs (bp). As we saw above, the simplest prokaryote, *Mycoplasma*, has several hundred protein-coding genes in a genome of 0.5 million bp. A rice plant, in contrast, has 37,544 genes.

- Eukaryotic genomes have more regulatory sequences—and many more regulatory proteins—than prokaryotic genomes. The greater complexity of eukaryotes requires much more regulation, which is evident in the many points of control associated with the expression of eukaryotic genes (see Figure 16.13).
- Much of eukaryotic DNA is noncoding. Distributed throughout many eukaryotic genomes are various kinds of DNA sequences that are not transcribed into mRNA, most notably introns and gene control sequences. As we discuss in Chapter 16, some noncoding sequences are transcribed into microRNAs. In addition, eukaryotic genomes contain various kinds of repeated sequences. These features are rare in prokaryotes.
- Eukaryotes have multiple chromosomes. The genomic "encyclopedia" of a eukaryote is separated into multiple "volumes." Each chromosome must have, at a minimum, three defining DNA sequences that we have described in previous chapters: an origin of replication (*ori*) that is recognized by the DNA replication machinery; a centromere region that holds the replicated chromosomes together before mitosis; and a telomeric sequence at each end of the chromosome that maintains chromosome integrity.

Model organisms reveal many characteristics of eukaryotic genomes

Most of the lessons learned from eukaryotic genomes have come from several simple model organisms that have been studied extensively: the yeast *Saccharomyces cerevisiae*, the nematode (roundworm) *Caenorhabditis elegans*, the fruit fly *Drosophila melanogaster*, and—representing plants—the thale cress, *Arabidopsis thaliana*. Model organisms have been chosen because they are relatively easy to grow and study in a laboratory, their genetics are well studied, and they exhibit characteristics that represent a larger group of organisms.

YEAST: THE BASIC EUKARYOTIC MODEL Yeasts are single-celled eukaryotes. Like most eukaryotes, they have membrane-enclosed organelles, such as the nucleus and endoplasmic reticulum, and a life cycle that alternates between haploid and diploid generations (see Figure 11.15).

TABLE 17.1

Representative Sequenced Genomes

ORGANISM	HAPLOID GENOME SIZE (Mb)	NUMBER OF GENES	PROTEIN-CODING SEQUENCE
Bacteria			
<i>M. genitalium</i>	0.58	485	88%
<i>H. influenzae</i>	1.8	1,738	89%
<i>E. coli</i>	4.6	4,377	88%
Yeasts			
<i>S. cerevisiae</i>	12.5	5,770	70%
<i>S. pombe</i>	12.5	4,929	60%
Plants			
<i>A. thaliana</i>	115	28,000	25%
Rice	390	37,544	12%
Animals			
<i>C. elegans</i>	100	19,427	25%
<i>D. melanogaster</i>	123	13,379	13%
Pufferfish	342	27,918	10%
Chicken	1,130	25,000	3%
Human	3,300	24,000	1.2%

Mb = millions of base pairs

While the prokaryote *E. coli* has a single circular chromosome with about 4.6 million bp and 4,290 protein-coding genes, budding yeast (*Saccharomyces cerevisiae*) has 16 linear chromosomes and a haploid content of more than 12.5 million bp, with 5,770 protein-coding genes. Gene inactivation studies similar to those carried out for *M. genitalium* (see Figure 17.7) indicate that fewer than 20 percent of these genes are essential to survival.

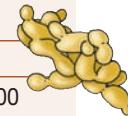
The most striking difference between the yeast genome and that of *E. coli* is in the number of genes for targeting proteins to organelles (Table 17.2). Both of these single-celled organisms appear to use about the same numbers of genes to perform the basic functions of cell survival. It is the compartmentalization of the eukaryotic yeast cell into organelles that requires it to have many more genes. This finding is direct, quantitative confirmation of something we have known for a century: the eukaryotic cell is structurally more complex than the prokaryotic cell.

THE NEMATODE: UNDERSTANDING EUKARYOTIC DEVELOPMENT In 1965 Sydney Brenner, fresh from being part of the team that first isolated mRNA, looked for a simple organism in which to study multicellularity. He settled on *Caenorhabditis elegans*, a millimeter-long nematode (roundworm) that normally lives in the soil. It can also live in the laboratory, where it has become a favorite

TABLE 17.2

Comparison of the Genomes of *E. coli* and Yeast

	E. COLI	YEAST
Genome length (base pairs)	4,640,000	12,068,000
Number of protein-coding genes	4,290	5,770
Proteins with roles in:		
Metabolism	650	650
Energy production/storage	240	175
Membrane transport	280	250
DNA replication/repair/recombination	120	175
Transcription	230	400
Translation	180	350
Protein targeting/secretion	35	430
Cell structure	180	250



model organism of developmental biologists (see Section 19.4). The nematode has a transparent body that develops over 3 days from a fertilized egg to an adult worm made up of nearly 1,000 cells. In spite of its small number of cells, the nematode has a nervous system, digests food, reproduces sexually, and ages. So it is not surprising that an intense effort was made to sequence the genome of this model organism.

The *C. elegans* genome (100 million bp) is eight times larger than that of yeast and has 3.5 times as many protein-coding genes (19,427). Gene inactivation studies have shown that the worm can survive in laboratory cultures with only 10 percent of these genes. So the “minimum genome” of a worm is about twice the size of that of yeast, which in turn is four times the size of the minimum genome for *Mycoplasma*. What do these extra genes do?

All cells must have genes for survival, growth, and division. In addition, the cells of multicellular organisms must have genes for holding cells together to form tissues, for cell differentiation, and for intercellular communication. Looking at Table 17.3, you will recognize functions that we discussed in earlier chapters,

TABLE 17.3

C. elegans Genes Essential to Multicellularity

FUNCTION	PROTEIN/DOMAIN	NUMBER OF GENES
Transcription control	Zinc finger; homeobox	540
RNA processing	RNA binding domains	100
Nerve impulse transmission	Gated ion channels	80
Tissue formation	Collagens	170
Cell interactions	Extracellular domains; glycosyltransferases	330
Cell-cell signaling	G protein-linked receptors; protein kinases; protein phosphatases	1,290



including gene regulation (see Chapter 16) and cell communication (see Chapter 7).

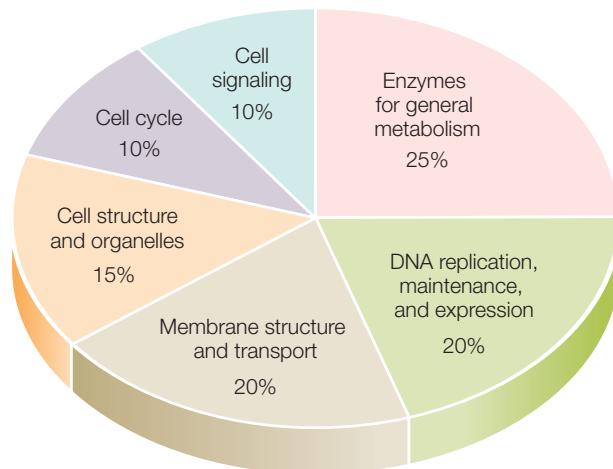
DROSOPHILA MELANOGASTER: RELATING GENETICS TO GENOMICS

The fruit fly *Drosophila melanogaster* is a famous model organism. Studies of fruit fly genetics resulted in the formulation of many basic principles of genetics (see Section 12.4). Over 2,500 mutations of *D. melanogaster* had been described by the 1990s when genome sequencing began, and this fact alone was a good reason for sequencing the fruit fly's DNA. The fruit fly is a much larger organism than *C. elegans*, both in size (it has 10 times more cells) and complexity, and it undergoes complicated developmental transformations from egg to larva to pupa to adult.

Not surprisingly, the fly's genome (about 123 million bp) is larger than that of *C. elegans*. But as we mentioned earlier, genome size does not necessarily correlate with the number of genes encoded. In this case, the larger fruit fly genome contains fewer genes (13,379) than the smaller nematode genome. **Figure 17.9** summarizes the functions of the *Drosophila* genes that have been characterized so far; this distribution is typical of complex eukaryotes.

ARABIDOPSIS: STUDYING THE GENOMES OF PLANTS About 250,000 species of flowering plants dominate the land and fresh water. But in the context of the history of life, the flowering plants are fairly young, having evolved only about 200 million years ago. The genomes of some plants are huge—for example, the genome of corn is about 3 billion bp, and that of wheat is 16 billion bp. So although we are naturally most interested in the genomes of plants we use as food and fiber, it is not surprising that scientists first chose to sequence a simpler flowering plant.

Arabidopsis thaliana, thale cress, is a member of the mustard family and has long been a favorite model organism of plant biologists. It is small (hundreds could grow and reproduce in the space occupied by this page) and easy to manipulate, and has a relatively small (115 million bp) genome.



17.9 Functions of the Eukaryotic Genome The distribution of gene functions in *Drosophila melanogaster* shows a pattern that is typical of many complex organisms.

TABLE 17.4

Arabidopsis Genes Unique to Plants

FUNCTION	NUMBER OF GENES
Cell wall and growth	42
Water channels	300
Photosynthesis	139
Defense and metabolism	94



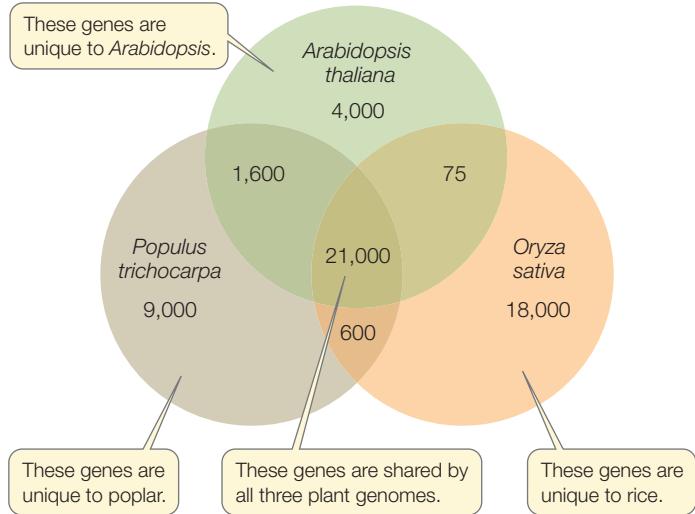
The *Arabidopsis* genome has about 28,000 protein-coding genes but, remarkably, many of these genes are duplicates and probably originated by chromosomal rearrangements. When these duplicate genes are subtracted from the total, about 15,000 unique genes are left—similar to the gene numbers found in fruit flies and nematodes. Indeed, many of the genes found in these animals have homologs (genes with very similar sequences) in *Arabidopsis* and other plants, suggesting that plants and animals have a common ancestor.

But *Arabidopsis* has some genes that distinguish it as a plant (**Table 17.4**). These include genes involved in photosynthesis, in the transport of water into the root and throughout the plant, in the assembly of the cell wall, in the uptake and metabolism of inorganic substances from the environment, and in the synthesis of specific molecules used for defense against microbes and herbivores (organisms that eat plants). These plant defense molecules may be a major reason why the number of protein-coding genes in plants is higher than in animals. Plants cannot escape their enemies or other adverse conditions as animals can, and so they must cope with the situation where they are. So they make tens of thousands of molecules to fight their enemies and adapt to the environment (see Chapter 39).

These plant-specific genes are also found in the genomes of other plants, including rice, the first major crop plant whose sequence has been determined. Rice (*Oryza sativa*) is the world's most important crop; it is a staple in the diet of 3 billion people. The larger genome in rice has a set of genes remarkably similar to that of *Arabidopsis*. More recently the genome of the poplar tree, *Populus trichocarpa*, was sequenced to gain insight into the potential for this rapidly growing tree to be used as a source of fixed carbon for making fuel. A comparison of the three genomes shows many genes in common, comprising the *basic plant genome* (**Figure 17.10**).

Eukaryotes have gene families

About half of all eukaryotic protein-coding genes exist as only one copy in the haploid genome (two copies in somatic cells). The rest are present in multiple copies, which arose from gene duplications. Over evolutionary time, different copies of genes have undergone separate mutations, giving rise to groups of closely related genes called **gene families**. Some gene families, such as those encoding the globin proteins that make up hemoglobin, contain only a few members; other families, such as the genes encoding the immunoglobulins that make up antibodies, have hundreds of members. In the human genome,



17.10 Plant Genomes Three plant genomes share a common set of approximately 21,000 genes that appear to comprise the “minimal” plant genome.

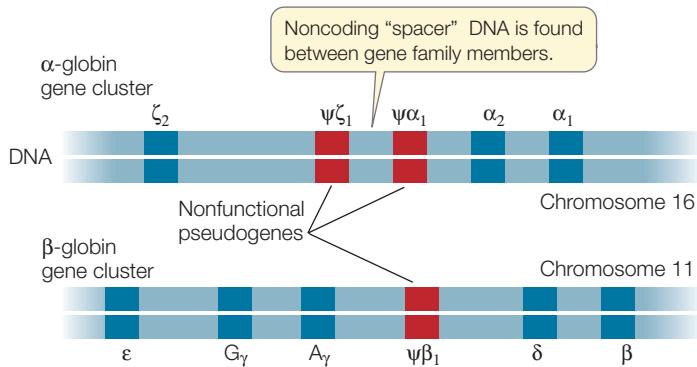
there are 24,000 protein-coding genes, but 16,000 distinct gene families. So only one-third of the human genes are unique.

The DNA sequences in a gene family are usually different from one another. As long as at least one member encodes a functional protein, the other members may mutate in ways that change the functions of the proteins they encode. For evolution, the availability of multiple copies of a gene allows for selection of mutations that provide advantages under certain circumstances. If a mutated gene is useful, it may be selected for in succeeding generations. If the mutated gene is a total loss, the functional copy is still there to carry out its role.

The gene family encoding the globins is a good example of the gene families found in vertebrates. These proteins are found in hemoglobin and myoglobin (an oxygen-binding protein present in muscle). The globin genes all arose long ago from a single common ancestral gene. In humans, there are three functional members of the α -globin cluster and five in the β -globin cluster (Figure 17.11). In adults, each hemoglobin molecule is a tetramer containing two identical α -globin subunits, two identical β -globin subunits, and four heme pigments (see Figure 3.10).

During human development, different members of the globin gene cluster are expressed at different times and in different tissues. This differential gene expression has great physiological significance. For example, hemoglobin containing γ -globin, a subunit found in the hemoglobin of the human fetus, binds O_2 more tightly than adult hemoglobin does. This specialized form of hemoglobin ensures that in the placenta, O_2 will be transferred from the mother’s blood to the developing fetus’s blood. Just before birth the liver stops synthesizing fetal hemoglobin and the bone marrow cells take over, making the adult forms (2α and 2β). Thus hemoglobins with different binding affinities for O_2 are provided at different stages of human development.

In addition to genes that encode proteins, many gene families include nonfunctional **pseudogenes**, which are designated with the Greek letter psi (Ψ) (see Figure 17.11). These pseudo-



17.11 The Globin Gene Family The α -globin and β -globin clusters of the human globin gene family are located on different chromosomes. The genes of each cluster are separated by noncoding “spacer” DNA. The nonfunctional pseudogenes are indicated by the Greek letter psi (Ψ). The γ gene has two variants, A_γ and G_γ .

genes result from mutations that cause a loss of function rather than an enhanced or new function. The DNA sequence of a pseudogene may not differ greatly from that of other family members. It may simply lack a promoter, for example, and thus fail to be transcribed. Or it may lack a recognition site needed for the removal of an intron, so that the transcript it makes is not correctly processed into a useful mature mRNA. In some gene families pseudogenes outnumber functional genes. Because some members of the family are functional, there appears to be little selection pressure for the deletion of pseudogenes.

Eukaryotic genomes contain many repetitive sequences

Eukaryotic genomes contain numerous repetitive DNA sequences that do not code for polypeptides. These include highly repetitive sequences, moderately repetitive sequences, and transposons.

Highly repetitive sequences are short (less than 100 bp) sequences that are repeated thousands of times in tandem (side-by-side) arrangements in the genome. They are not transcribed. Their proportion in eukaryotic genomes varies, from 10 percent in humans to about half the genome in some species of fruit flies. Often they are associated with heterochromatin, the densely packed, transcriptionally inactive part of the genome. Other highly repetitive sequences are scattered around the genome. For example, short tandem repeats (STRs) of 1–5 bp can be repeated up to 100 times at a particular chromosomal location. The copy number of an STR at a particular location varies between individuals and is inherited. In Chapter 15 we describe how STRs can be used in the identification of individuals (DNA fingerprinting).

Moderately repetitive sequences are repeated 10–1000 times in the eukaryotic genome. These sequences include the genes that are transcribed to produce tRNAs and rRNAs, which are used in protein synthesis. The cell makes tRNAs and rRNAs constantly, but even at the maximum rate of transcription, single copies of the tRNA and rRNA genes would be inadequate to supply the large amounts of these molecules needed by most cells. Thus the genome has multiple copies of these genes.

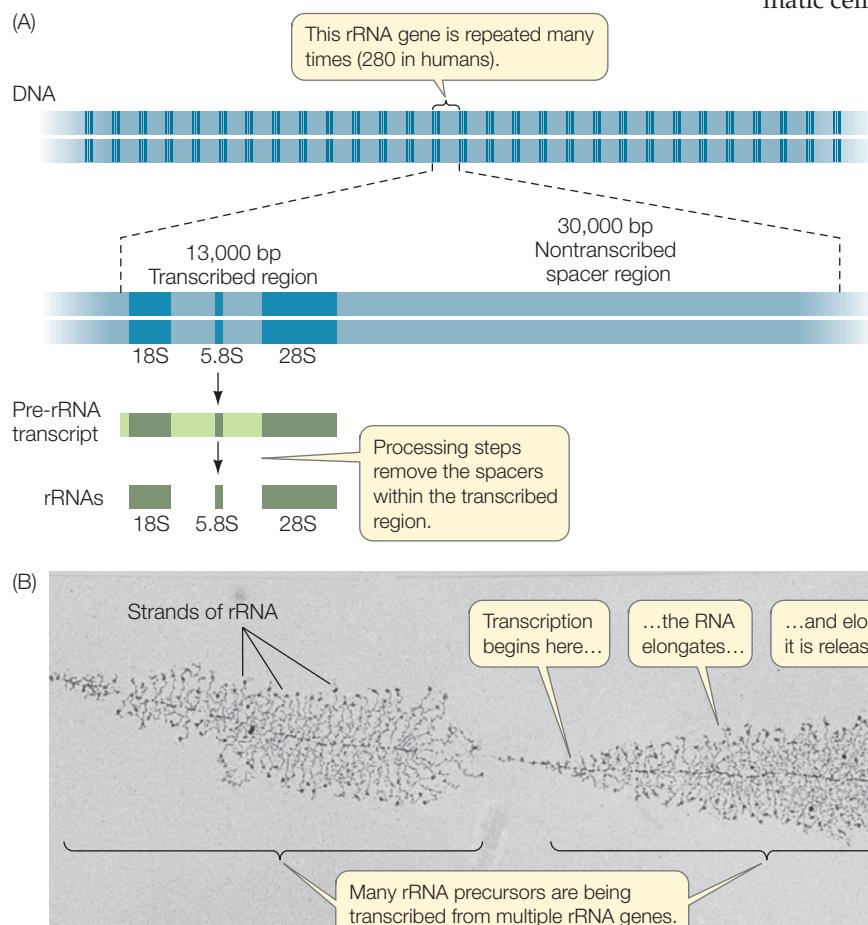
In mammals, four different rRNA molecules make up the ribosome: the 18S, 5.8S, 28S, and 5S rRNAs. (The S stands for Svedberg unit, which is a measure of size.) The 18S, 5.8S, and

28S rRNAs are transcribed together as a single precursor RNA molecule (Figure 17.12). As a result of several posttranscriptional steps, the precursor is cut into the final three rRNA products, and the noncoding “spacer” RNA is discarded. The sequence encoding these RNAs is moderately repetitive in humans: a total of 280 copies of the sequence are located in clusters on five different chromosomes.

TRANSPOSONS Apart from the RNA genes, most moderately repetitive sequences are not stably integrated into the genome. Instead, these sequences can move from place to place, and are thus called transposable elements or transposons. Prokaryotes also have transposons (see Figure 17.6). Transposons make up over 40 percent of the human genome and about 50 percent of the maize genome, although the percentage is smaller (3–10 percent) in many other eukaryotes.

There are four main types of transposons in eukaryotes:

1. **SINEs** (short interspersed elements) are up to 500 bp long and are transcribed but not translated. There are about 1.5 million of them scattered over the human genome, making up about 15 percent of the total DNA content. A single type, the 300-bp Alu element, accounts for 11 percent of the human genome; it is present in a million copies.
2. **LINEs** (long interspersed elements) are up to 7,000 bp long, and some are transcribed and translated into proteins. They constitute about 17 percent of the human genome.



SINEs and LINEs move about the genome in a distinctive way: they are transcribed into RNA, which then acts as a template for new DNA. The new DNA becomes inserted at a new location in the genome. This “copy and paste” mechanism results in two copies of the transposon: one at the original location and the other at a new location.

3. **Retrotransposons** also make RNA copies of themselves when they move about the genome. Some of them encode proteins needed for their own transposition, and others do not. SINEs and LINEs are types of retrotransposons. Non-SINE, non-LINE retrotransposons constitute about 8 percent of the human genome.
4. **DNA transposons** do not use RNA intermediates. Like some prokaryotic transposable elements, they are excised from the original location and become inserted at a new location without being replicated.

What role do these moving sequences play in the cell? The best answer so far seems to be that transposons are simply cellular parasites that can be replicated. The insertion of a transposon at a new location can have important consequences. For example, the insertion of a transposon into the coding region of a gene results in a mutation (see Figure 17.8). This phenomenon accounts for a few rare forms of several genetic diseases in humans, including hemophilia and muscular dystrophy. If the insertion of a transposon takes place in the germ line, a gamete with a new mutation results. If the insertion takes place in a somatic cell, cancer may result.

Sometimes an adjacent gene can be replicated along with a transposon, resulting in a gene duplication. A transposon can carry a gene, or a part of it, to a new location in the genome, shuffling the genetic material and creating new genes. Clearly, transposition stirs the genetic pot in the eukaryotic genome and thus contributes to genetic variation.

Section 5.5 describes the theory of *endosymbiosis*, which proposes that chloroplasts and mitochondria are the descendants of once free-living prokaryotes. Transposons may have played a role in endosymbiosis. In living eukaryotes the chloroplasts and mitochondria contain some DNA, but the nucleus contains most of the genes

17.12 A Moderately Repetitive Sequence Codes for rRNA (A) This rRNA gene, along with its nontranscribed spacer region, is repeated 280 times in the human genome, with clusters on five chromosomes. (B) This electron micrograph shows transcription of multiple rRNA genes.

17.13 Sequences in the Eukaryotic Genome There are many types of DNA sequences. Some are transcribed, and some of those sequences are translated.

that encode the organelles' proteins. If the organelles were once independent, they must originally have contained all of those genes. How did the genes move to the nucleus? They may have done so by DNA transpositions between organelles and the nucleus, which still occur today. The DNA that remains in the organelles may be the remnants of more complete prokaryotic genomes.

See **Figure 17.13** for a summary of the various types of sequences in the human genome.

17.3 RECAP

The sequencing of the genomes of model organisms demonstrated common features of the eukaryotic genome, including the presence of repetitive sequences and transposons. Some eukaryotic genes are in families, which may include members that are mutated and nonfunctional. Some sequences are transcribed, but others are not.

- What are the major differences between prokaryotic and eukaryotic genomes? **See p. 375**
- Describe one function of genes found in *C. elegans* that has no counterpart in the genome of yeast. **See p. 376 and Table 17.3**
- What is the evolutionary role of eukaryotic gene families? **See p. 377**
- Why are there multiple copies of sequences coding for rRNA in the mammalian genome? **See p. 378**
- What effects can transposons have on a genome? **See p. 379**

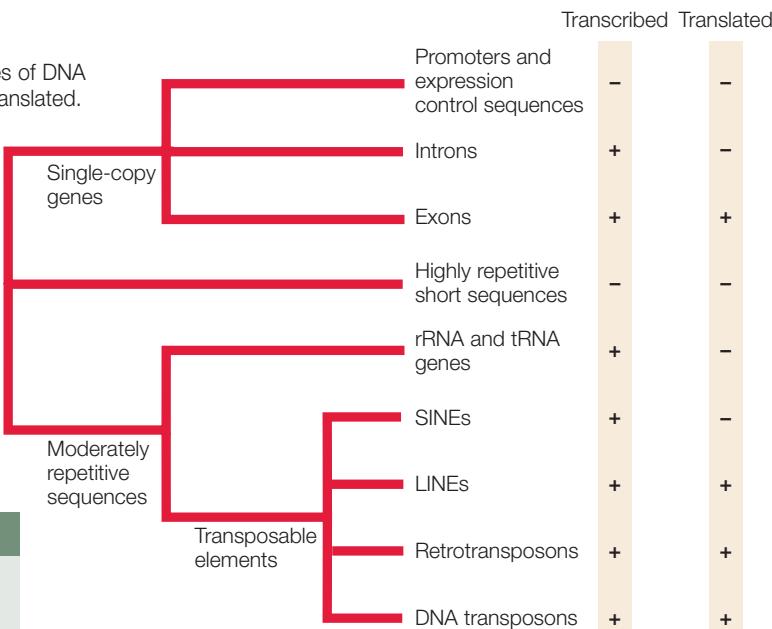
The analysis of eukaryotic genomes has resulted in an enormous amount of useful information, as we have seen. In the next section we look more closely at the human genome.

17.4 What Are the Characteristics of the Human Genome?

By the start of 2005 the first human genome sequences were completed, two years ahead of schedule and well under budget. The published sequences, one produced by the publicly funded Human Genome Project, and the other by a private company, were haploid genomes that were composites of several people. Since 2005, the diploid genomes of several individuals have been sequenced and published.

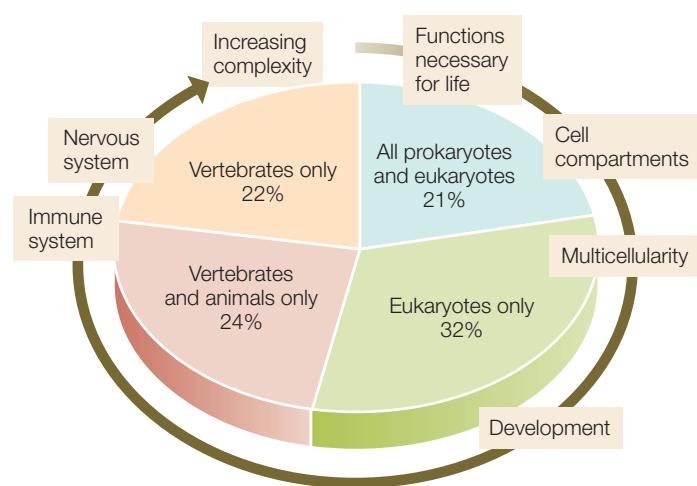
The human genome sequence held some surprises

The following are just some of the interesting facts that we have learned about the human genome:



- Of the 3.3 billion base pairs in the haploid human genome, fewer than 2 percent (about 24,000 genes) make up protein-coding regions. This was a surprise. Before sequencing began, humans were estimated, based on the diversity of their proteins, to have 80,000–150,000 genes. The actual number of genes—not many more than in a fruit fly—means that posttranscriptional mechanisms (such as alternative splicing) must account for the observed number of proteins in humans. That is, the average human gene must code for several different proteins.
- The average gene has 27,000 base pairs. Gene sizes vary greatly, from about 1,000 to 2.4 million base pairs. Variation in gene size is to be expected given that human proteins (and RNAs) vary in size, from 100 to about 5,000 amino acids per polypeptide chain.
- Virtually all human genes have many introns.
- Over 50 percent of the genome is made up of transposons and other highly repetitive sequences. Repetitive sequences near genes are GC-rich, while those farther away from genes are AT-rich.
- Most of the genome (about 97 percent) is the same in all people. Despite this apparent homogeneity, there are, of course, many individual differences. Scientists have mapped over 7 million single nucleotide polymorphisms (SNPs) in humans.
- Genes are not evenly distributed over the genome. Chromosome 19 is packed densely with genes, while chromosome 8 has long stretches without coding regions. The Y chromosome has the fewest genes (231), while chromosome 1 has the most (2,968).

Comparisons between sequenced genomes from prokaryotes and eukaryotes have revealed some of the evolutionary relationships between genes. Some genes are present in both prokaryotes and eukaryotes; others are only in eukaryotes; still others are only in animals, or only in vertebrates (**Figure 17.14**).



17.14 Evolution of the Genome A comparison of the human and other genomes has revealed how genes with new functions have been added over the course of evolution. Each percentage number refers to genes in the human genome. Thus, 21 percent of human genes have homologs in prokaryotes and other eukaryotes, 32 percent of human genes occur only in other eukaryotes, and so on.

More comparative genomics is possible now that the genomes of two other primates, the chimpanzee and the rhesus macaque, have been sequenced. The chimpanzee is evolutionarily close to humans, and shares 95 percent of the human genome sequence. The more distantly related rhesus macaque shares 91 percent of the human sequence. The search is on for a set of human genes that differ from the other primates and “make humans human.”

Human genomics has potential benefits in medicine

Complex phenotypes are determined not by single genes, but by multiple genes interacting with the environment. The single-allele models of phenylketonuria and sickle-cell anemia (see Chapter 15) do not apply to such common disorders as diabetes, heart disease, and Alzheimer’s disease. To understand the genetic bases of these diseases, biologists are now using rapid genotyping technologies to create “haplotype maps,” which are used to identify SNPs (pronounced “snips”) that are linked to genes involved in disease.

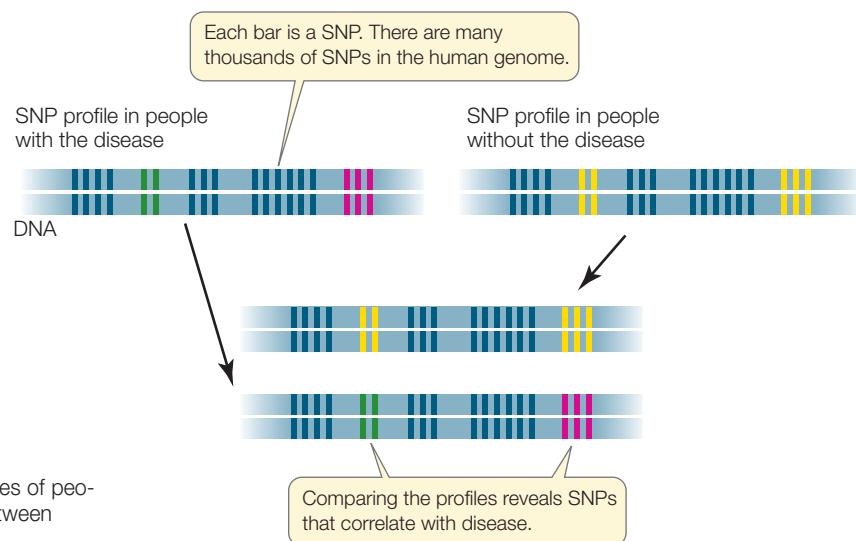
HAPLOTYPE MAPPING The SNPs that differ between individuals are not inherited as independent alleles. Rather, *a set of SNPs that are present on a segment of chromosome are usually inherited as a unit*. This linked piece of a chromosome is called a **haplotype**. You can think of the haplotype as a sentence and the SNP as

a word in the sentence. Analyses of haplotypes in humans from all over the world have shown that there are at most 500,000 common variations.

GENOTYPING TECHNOLOGY AND PERSONAL GENOMICS New technologies are continually being developed to analyze thousands or millions of SNPs in the genomes of individuals. Such technologies include rapid sequencing methods and DNA microarrays that depend on DNA hybridization to identify specific SNPs. For example, a microarray of 500,000 SNPs has been used to analyze thousands of people to find out which SNPs are associated with specific diseases. The aim is to *correlate the SNP-defined haplotype with a disease state*. The amount of data is prodigious: 500,000 SNPs, thousands of people, thousands of medical records. With so much natural variation, statistical measures of association between a haplotype and a disease need to be very rigorous.

These association tests have revealed particular haplotypes or alleles that are associated with modestly increased risks for such diseases as breast cancer, diabetes, arthritis, obesity, and coronary heart disease (Figure 17.15 and Table 17.5). Private companies will now scan a human genome for these variants—and the price for this service keeps getting lower. However, at this point it is unclear what a person without symptoms should do with the information, since multiple genes, environmental influences, and epigenetic effects all contribute to the development of these diseases.

Of course, the best way to analyze a person’s genome is by actually sequencing it. Until recently, this was prohibitively expensive. As we mentioned earlier, DNA pioneer James Watson’s genome cost over \$1 million, certainly too much for a typical person or insurance company to afford in the context of health care. But with advances in sequencing technologies the cost is decreasing rapidly. One new method automatically sequences protein-coding exons only, for example. Once the cost of genome sequencing is within an affordable range, SNP testing will be superseded.

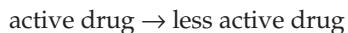


17.15 SNP Genotyping and Disease Scanning the genomes of people with and without particular diseases reveals correlations between SNPs and complex diseases.

TABLE 17.5
SNP Human Genome Scans and Diseases

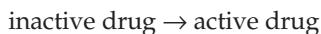
DISEASE	LOCATION OF SNP (CHROMOSOME NUMBER)	% INCREASED RISK	
		HETEROZYGOTES	HOMOZYGOTES
Breast cancer	8	20	63
Coronary heart disease	9	20	56
Heart attack	9	25	64
Obesity	16	32	67
Diabetes	10	65	277
Prostate cancer	8	26	58

PHARMACOGENOMICS Genetic variation can affect how an individual responds to a particular drug. For example, a drug may be chemically modified in the liver to make it more or less active. Consider an enzyme that catalyzes the following reaction:

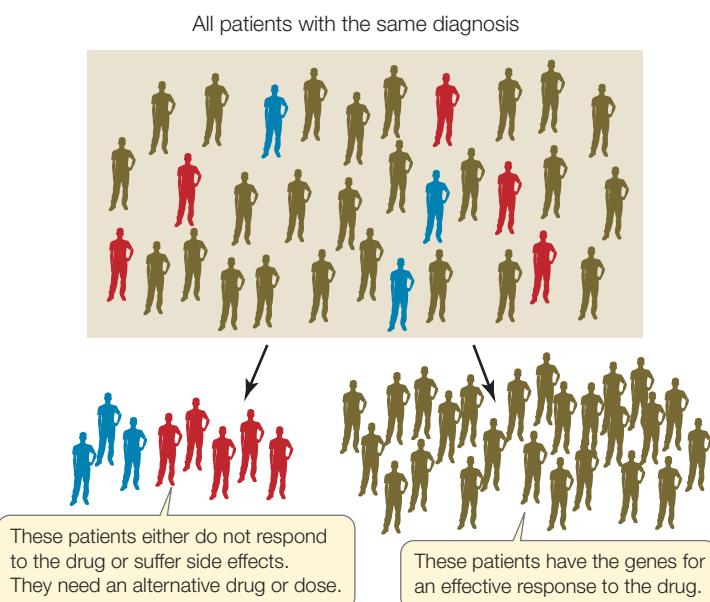


A mutation in the gene that encodes this enzyme may make the enzyme less active. For a given dose of the drug, a person with the mutation would have more active drug in the bloodstream than a person without the mutation. So the effective dose of the drug would be lower in these people.

Now consider a different case, in which the liver enzyme is needed to make the drug active:



A person carrying a mutation in the gene encoding this liver enzyme would not be affected by the drug, since the activating enzyme is not present.



17.16 Pharmacogenomics Correlations between genotypes and responses to drugs will help physicians develop personalized medical care.

The study of how an individual's genome affects his or her response to drugs or other outside agents is called **pharmacogenomics**. This type of analysis makes it possible to predict whether a drug will be effective. The objective is to *personalize drug treatment* so that a physician can know in advance whether an individual will benefit from a particular drug (Figure 17.16). This approach might also be used to reduce the incidence of adverse drug reactions by identifying individuals that will metabolize a drug slowly, which can lead to a dangerously high level of the drug in the body.

17.4 RECAP

The haploid human genome has 3.3 billion base pairs, but less than 2 percent of the genome codes for proteins. Most human genes are subject to alternative splicing; this may account for the fact that there are more proteins than genes. SNP mapping to find correlations with disease and drug susceptibility holds promise for personalized medicine.

- What are some of the major characteristics of the human genome? See p. 380
- How does SNP mapping work in personalized medicine? See pp. 381–382 and Figures 17.15 and 17.16

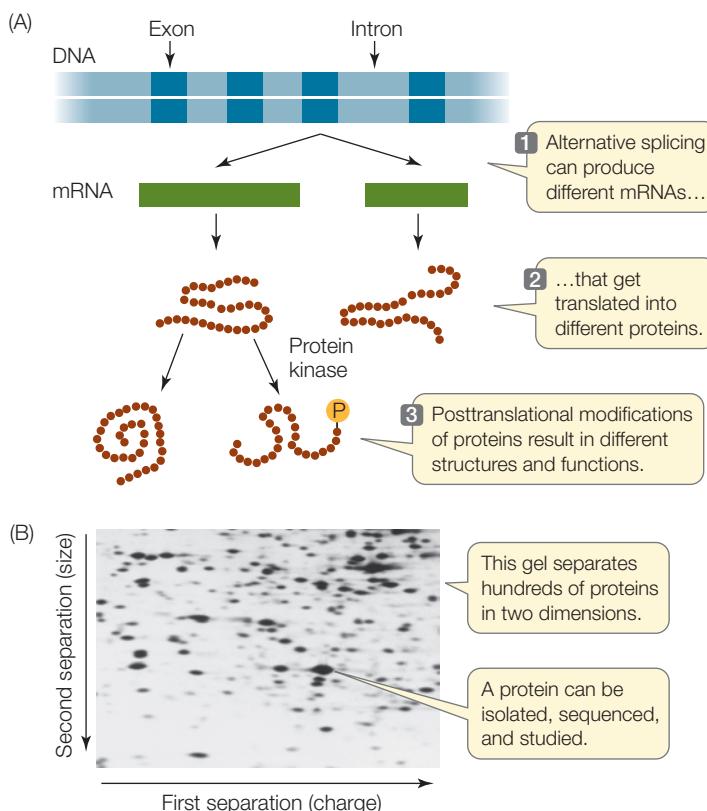
Genome sequencing has had great success in advancing biological understanding. High-throughput technologies are now being applied to other components of the cell: proteins and metabolites. We now turn to the results of these studies.

17.5 What Do the New Disciplines Proteomics and Metabolomics Reveal?

“The human genome is the book of life.” Statements like this were common at the time the human genome sequence was first revealed. They reflect “genetic determinism,” that a person’s phenotype is determined by his or her genotype. But is an organism just a product of gene expression? We know that it is not. The proteins and small molecules present in any cell at a given point in time reflect not just gene expression but modifications by the intracellular and extracellular environment. Two new fields have emerged to complement genomics and take a more complete snapshot of a cell and organism—proteomics and metabolomics.

The proteome is more complex than the genome

As mentioned above, many genes encode more than a single protein (Figure 17.17A). Alternative splicing leads to different combinations of exons in the mature mRNAs transcribed from a single gene (see Figure 16.22). Posttranslational modifications also increase the number of proteins that can be derived from one gene (see Figure 14.22). The **proteome** is the sum total of the proteins produced by an organism, and it is more complex than its genome.



17.17 Proteomics (A) A single gene can code for multiple proteins. (B) A cell's proteins can be separated on the basis of charge and size by two-dimensional gel electrophoresis. The two separations can distinguish most proteins from one another.

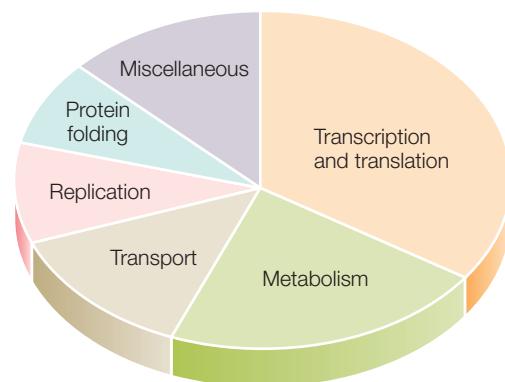
Two methods are commonly used to analyze the proteome:

- Because of their unique amino acid compositions (primary structures), most proteins have unique combinations of electric charge and size. On the basis of these two properties, they can be separated by two-dimensional gel electrophoresis. Thus isolated, individual proteins can be analyzed, sequenced, and studied (**Figure 17.17B**).
- Mass spectrometry uses electromagnets to identify proteins by the masses of their atoms and displays them as peaks on a graph.

The ultimate aim of proteomics is just as ambitious as that of genomics. While genomics seeks to describe the genome and its expression, proteomics seeks to identify and characterize all of the expressed proteins.

Comparisons of the proteomes of humans and other eukaryotic organisms have revealed a common set of proteins that can be categorized into groups with similar amino acid sequences and similar functions. Forty-six percent of the yeast proteome, 43 percent of the worm proteome, and 61 percent of the fly proteome are shared by the human proteome. Functional analyses indicate that this set of 1,300 proteins provide the basic metabolic functions of a eukaryotic cell, such as glycolysis, the citric acid cycle, membrane transport, protein synthesis, DNA replication, and so on. (**Figure 17.18**).

Of course, these are not the only human proteins. There are many more, which presumably distinguish us as *human* eukary-



17.18 Proteins of the Eukaryotic Proteome About 1,300 proteins are common to all eukaryotes and fall into these categories. Although their amino acid sequences may differ to a limited extent, they perform the same essential functions in all eukaryotes.

otic organisms. As we have mentioned before, proteins have different functional regions called domains (for example, a domain for binding a substrate, or a domain for spanning a membrane). While a particular organism may have many unique proteins, those proteins are often just unique combinations of domains that exist in other organisms. *This reshuffling of the genetic deck is a key to evolution.*

Metabolomics is the study of chemical phenotype

Studying genes and proteins gives a limited picture of what is going on in a cell. But as we have seen, both gene function and protein function are affected by the internal and external environments of the cell. Many proteins are enzymes and their activities affect the concentrations of their substrates and products. So as the proteome changes, so will the abundances of these often-small molecules, called metabolites. The **metabolome** is the quantitative description of all of the *small molecules* in a cell or organism. These include:

- Primary metabolites* involved in normal processes, such as intermediates in pathways like glycolysis. This category also includes hormones and other signaling molecules.
- Secondary metabolites*, which are often unique to particular organisms or groups of organisms. They are often involved in special responses to the environment. Examples are antibiotics made by microbes, and the many chemicals made by plants that are used in defense against pathogens and herbivores.

Not surprisingly, measuring metabolites involves sophisticated analytical instruments. If you have studied organic or analytical chemistry, you may be familiar with gas chromatography and high-performance liquid chromatography, which separate molecules, and mass spectrometry and nuclear magnetic resonance spectroscopy, which are used to identify them. These measurements result in “chemical snapshots” of cells or organisms, which can be related to physiological states.

There has been some progress in defining the human metabolome. A database created by David Wishart and col-

leagues at the University of Alberta contains over 6,500 metabolite entries. The challenge now is to relate levels of these substances to physiology. For example, you probably know high levels of glucose in the blood are associated with diabetes. But what about early stages of heart disease? There may be a pattern of metabolites that is diagnostic of this disease. This could aid in early diagnosis and treatment.

Plant biologists are far ahead of medical researchers in the study of metabolomics. Over the years, tens of thousands of secondary metabolites have been identified in plants, many of them made in response to environmental challenges. Some of these are discussed in Chapter 39. The metabolome of the model organism *Arabidopsis thaliana* is being described, and will give insight into how a plant copes with stresses such as drought or pathogen attack. This knowledge could be helpful in optimizing plant growth for agriculture.

17.5 RECAP

The proteome is the total of all proteins produced by an organism. There are more proteins than genes in the genome. The metabolome is the total content of small molecules such as intermediates in metabolism, hormones, and secondary metabolites. The proteome and the metabolome can be analyzed using chemical methods that separate and identify molecules.

- How is the proteome analyzed? [See p. 383 and Figure 17.17](#)
- Explain the differences between genome, proteome, and metabolome.

CHAPTER SUMMARY

17.1 How Are Genomes Sequenced?

- The sequencing of genomes required the development of ways to cut large chromosomes into fragments, sequence each of the fragments, and then line them up on the chromosome. [Review Figure 17.1, ANIMATED TUTORIAL 17.1](#)
- **Hierarchical sequencing** involves mapping the genome with genetic markers, cutting the genome into smaller pieces and sequencing them, then lining up the sequences using the markers.
- **Shotgun sequencing** involves directly cutting the genome into overlapping fragments, sequencing them, and using a computer to line up the sequences.
- DNA sequencing technologies involve labeled nucleotides that terminate the growing polynucleotide chain. [Review Figure 17.2](#)
- Rapid, automated methods for **high-throughput sequencing** are being developed. [Review Figure 17.3, ANIMATED TUTORIAL 17.2](#)

17.2 What Have We Learned from Sequencing Prokaryotic Genomes?

- DNA sequencing is used to study the genomes of prokaryotes that are important to humans and ecosystems.
- **Functional genomics** aims to determine the functions of gene products. **Comparative genomics** involves comparisons of genes and genomes from different organisms to identify common features and functions.
- **Transposable elements** and **transposons** can move about the genome. [Review Figure 17.6](#)
- **Metagenomics** is the identification of DNA sequences without first isolating, growing and identifying the organisms present in an environmental sample. Many of these sequences are from prokaryotes that were heretofore unknown to biologists. [Review Figure 17.7](#)
- Transposon mutagenesis can be used to inactivate genes one by one. Then the organism can be tested for survival. In this way, a minimal genome of less than 350 genes was identified for the bacterium *Mycoplasma genitalium*. [Review Figure 17.8](#)

17.3 What Have We Learned from Sequencing Eukaryotic Genomes?

- Genome sequences from model organisms have demonstrated some common features of the eukaryotic genome. In addition, there are specialized genes for cellular compartmentation, development, and features unique to plants. [Review Tables 17.1–17.4 and Figures 17.9 and 17.10](#)
- Some eukaryotic genes exist as members of **gene families**. Proteins may be made from these closely related genes at different times and in different tissues. Some members of gene families may be nonfunctional **pseudogenes**.
- Repeated sequences are present in the eukaryotic genome.
- **Moderately repeated sequences** include those coding for rRNA. [Review Figure 17.12](#)

17.4 What Are the Characteristics of the Human Genome?

- The haploid human genome has 3.3 billion base pairs.
- Only 2 percent of the genome codes for proteins; the rest consists of repeated sequences and noncoding DNA.
- Virtually all human genes have introns, and alternative splicing leads to the production of more than one protein per gene.
- SNP genotyping correlates variations in the genome with diseases or drug sensitivity. It may lead to personalized medicine. [Review Figure 17.15](#)
- **Pharmacogenomics** is the analysis of genetics as applied to drug metabolism.

17.5 What Do the New Disciplines of Proteomics and Metabolomics Reveal?

- The **proteome** is the total protein content of an organism.
- There are more proteins than protein-coding genes in the genome.
- The proteome can be analyzed using chemical methods that separate and identify proteins. These include two-dimensional electrophoresis and mass spectrometry. [See Figure 17.17](#)
- The **metabolome** is the total content of small molecules, such as intermediates in metabolism, hormones, and secondary metabolites.

SEE WEB ACTIVITY 17.1 for a concept review of this chapter.

SELF-QUIZ

1. Eukaryotic protein-coding genes differ from their prokaryotic counterparts in that eukaryotic genes
 - a. are double-stranded.
 - b. are present in only a single copy.
 - c. contain introns.
 - d. have promoters.
 - e. are transcribed into mRNA.
2. A comparison of the genomes of yeast and bacteria shows that only yeast has many genes for
 - a. energy metabolism.
 - b. cell wall synthesis.
 - c. intracellular protein targeting.
 - d. DNA-binding proteins.
 - e. RNA polymerase.
3. The genomes of the fruit fly and the nematode are similar to that of yeast, except that the former organisms have many genes for
 - a. intercellular signaling.
 - b. synthesis of polysaccharides.
 - c. cell cycle regulation.
 - d. intracellular protein targeting.
 - e. transposable elements.
4. The minimum genome of *Mycoplasma genitalium*
 - a. has 100 genes.
 - b. has been used to create new species.
 - c. has an RNA genome.
 - d. is larger than the genome of *E. coli*.
 - e. was derived by transposon mutagenesis.
5. Which is *not* true of metagenomics?
 - a. It has been done with bacteria.
 - b. It is done on rRNA sequences.
 - c. It has revealed many new metabolic capacities.
 - d. It involves extracting DNA from the environment.
 - e. It cannot be done on seawater.
6. Transposons
 - a. always use RNA for replication.
 - b. are approximately 50 bp long.
 - c. are made up of either DNA or RNA.
 - d. do not contain genes coding for proteins.
 - e. make up about 40 percent of the human genome.
7. Vertebrate gene families
 - a. have mostly inactive genes.
 - b. include the globins.
 - c. are not produced by gene duplications.
 - d. increase the number of unique genes in the genome.
 - e. are not transcribed.
8. The DNA sequences that code for eukaryotic rRNA
 - a. are transcribed only at the ribosome.
 - b. are repeated hundreds of times.
 - c. contain all the genes clustered directly beside one another.
 - d. are on only one human chromosome.
 - e. are identical to the sequences that code for miRNA.
9. The human genome
 - a. contains very few repeated sequences.
 - b. has 3.3 billion base pairs.
 - c. was sequenced by hierarchical sequencing only.
 - d. has genes evenly distributed along chromosomes.
 - e. has few genes with introns.
10. Which of the following about genome sequencing is true?
 - a. In hierarchical sequencing, but not high-throughput sequencing, DNA is amplified in BAC vectors.
 - b. In hierarchical sequencing, a genetic map is made after the DNA is sequenced.
 - c. Shotgun sequencing is considerably slower than hierarchical sequencing.
 - d. The human genome was first sequenced by high-throughput methods.
 - e. DNA sequence determination by chain termination is the basis of shotgun sequencing only.

FOR DISCUSSION

1. In rats, a protein-coding gene 1,440 bp long codes for an enzyme made up of 192 amino acids. Discuss this apparent discrepancy. How long would the initial and final mRNA transcripts be?
2. The genomes of rice, wheat, and corn are similar to one another and to that of *Arabidopsis* in many ways. Discuss how these plants might nevertheless have very different proteins.
3. Why are the proteome and the metabolome more complex than the genome?

ADDITIONAL INVESTIGATION

It is the year 2025. You are taking care of a patient who is concerned about having an early stage of kidney cancer. His mother died from this disease.

- a. Assume that the SNPs linked to genes involved in the development of this type of cancer have been identified. How would you determine if this man has a genetic predisposition for developing kidney cancer? Explain how you would do the analysis.

- b. How might you develop a metabolomic profile for kidney cancer and then use it to determine whether your patient has kidney cancer?
- c. If the patient was diagnosed with cancer by the methods in (a) and (b), how would you use pharmacogenomics to choose the right medications to treat the tumor in this patient?

18

Recombinant DNA and Biotechnology

Pollution fighters

In the summer of 1990, soldiers from Iraq invaded neighboring Kuwait. The reason was oil: the Iraqis were angry because Kuwait was pumping too much of it, keeping prices low. Six months later, a United Nations-sponsored coalition army from more than 30 countries drove the Iraqis out of Kuwait and back to their homeland. For Kuwait, the Gulf War was a success, but it left an environmental disaster. As they fled, the Iraqi soldiers set fire to more than 700 oil wells. It took over six months to put the fires out, and in the meantime an astounding 250 million gallons of crude oil were released into the desert. Twenty years later, much of the oil remains as a gooey coating, severely affecting the organisms that live there.

The government of Kuwait is using a variety of processes to get rid of the contaminating oil. Among them is the addition of bacteria that break down and consume the oil, utilizing the hydrocarbons in it as an energy

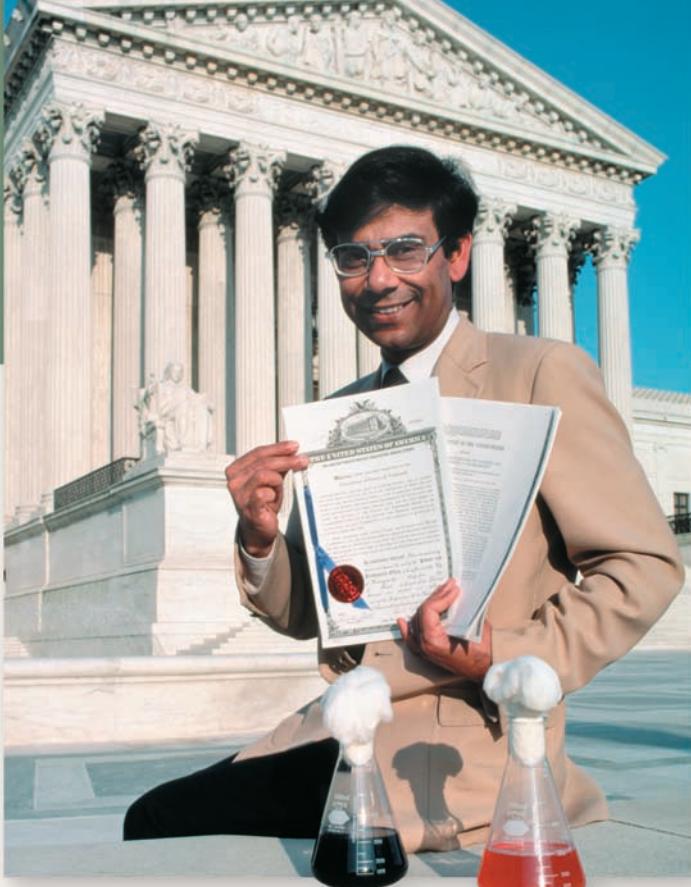
source for growth. This process—using an organism to remove a pollutant—is called bioremediation. The Kuwait episode is not the first major use of bacteria for bioremediation. In 1989, the oil tanker *Exxon Valdez* ran aground near the Alaskan shore, releasing 11 million gallons of crude oil along 500 miles of shoreline. Physical methods such as skimming the water were used to remove more than half of the oil. Nitrogen and phosphorus salts were then sprayed on the oily rocks to stimulate the growth of oil-consuming bacteria already there, and other bacteria were added as part of the recovery effort. The oil gradually disappeared.

Some species of bacteria, because of their genetic capacity to produce unusual enzymes and biochemical pathways, thrive on all sorts of nutrients besides the usual glucose, including pollutants. Scientists have discovered these organisms simply by mixing polluted soil with water and

seeing what grows. Many of the genes coding for enzymes involved in breaking down crude oil are carried on small chromosomes called plasmids. In 1971, Ananda Chakrabarty at the General Electric Research Center in New York used genetic crosses to develop a single strain of the bacterium *Pseudomonas* with multiple plasmids carrying genes for the breakdown of various hydrocarbons in oil. He and his company applied for a patent to legally protect their discovery and profit from it. In a landmark case, the U.S. Supreme Court ruled in 1980 that “a live, human-made microorganism is



The Spoils of War Massive oil spills occurred in Kuwait during the 1991 Gulf War.



Using Biotechnology to Clean Up the Environment

Ananda Chakrabarty received the first patent for a genetically modified organism, a bacterium that breaks down crude oil.

patentable” under the U.S. Constitution. Since then other bacteria have been patented that remove toxic metals such as mercury and copper from soils. In these cases, the bacteria use metabolic pathways to convert the metals to biologically inert forms.

The Supreme Court ruling came at a time when new laboratory methods were being developed to insert specific DNA sequences into organisms by recombinant DNA technology. Since then, an entirely new biotechnology industry has sprung up, its activities legally protected. The resulting flood of patents for DNA sequences and genetically modified organisms continues to this day.

IN THIS CHAPTER we will describe some of the techniques that are used to manipulate DNA. First, we will describe how DNA molecules are cut into smaller fragments and how these fragments are spliced together to create recombinant DNA. This will lead to a discussion of how recombinant DNA is introduced into suitable host cells. After describing some other ways to manipulate DNA, we will show how scientists have applied these methods to create a new biotechnology industry.

CHAPTER OUTLINE

- 18.1 What Is Recombinant DNA?**
- 18.2 How Are New Genes Inserted into Cells?**
- 18.3 What Sources of DNA Are Used in Cloning?**
- 18.4 What Other Tools Are Used to Study DNA Function?**
- 18.5 What Is Biotechnology?**
- 18.6 How Is Biotechnology Changing Medicine, Agriculture, and the Environment?**

18.1 What Is Recombinant DNA?

You are familiar with *restriction endonucleases* (restriction enzymes), which occur naturally in bacteria and are used in the laboratory to cut DNA into fragments (see Chapter 15). Our focus in Chapter 15 was on the use of these enzymes for detecting mutations. In this chapter we examine how they are used, along with other enzymes, to construct recombinant DNA.

During the late 1960s, scientists discovered other enzymes that act on DNA. One of these is **DNA ligase**, which catalyzes the joining of DNA fragments. This is the enzyme that joins Okazaki fragments during DNA replication (see Section 13.3). Once they had isolated restriction enzymes and DNA ligase, scientists could use these enzymes to cut DNA into fragments and then splice them together in new combinations. Stanley Cohen and Herbert Boyer did just that in 1973. They used restriction enzymes to cut sequences from two *E. coli* plasmids (small chromosomal DNAs—see Figure 12.27) containing different antibiotic resistance genes. Then they used DNA ligase to join the fragments together. The resulting plasmid, when inserted into new *E. coli* cells, gave those cells resistance to both antibiotics (Figure 18.1). The era of **recombinant DNA**—a DNA molecule made in the laboratory that is derived from at least two genetic sources—was born.

Hundreds of different restriction enzymes are now available. They recognize *palindromic* DNA sequences—sequences that read the same way in both directions. For example, you can read the DNA recognition sequence for the restriction enzyme *EcoR1* from 5' to 3' as GAATTC on both strands:



Some restriction enzymes cut the DNA straight through the middle of the palindrome, generating “blunt-ended” fragments. Others, such as *EcoRI*, make staggered cuts—they cut one strand of the double helix several bases away from where they cut the other (Figure 18.2). After *EcoRI* makes its two cuts in the complementary strands, the ends of the strands are held together only by the hydrogen bonds between four base pairs. These hydrogen bonds are too weak to persist at warm temperatures (above room temperature), so the DNA separates into fragments when it is warmed. As a result, each fragment carries a single-stranded “overhang” at the location of each cut. These overhangs are called **sticky ends** because they have specific base sequences that can bind by base pairing with complementary sticky ends.

INVESTIGATING LIFE

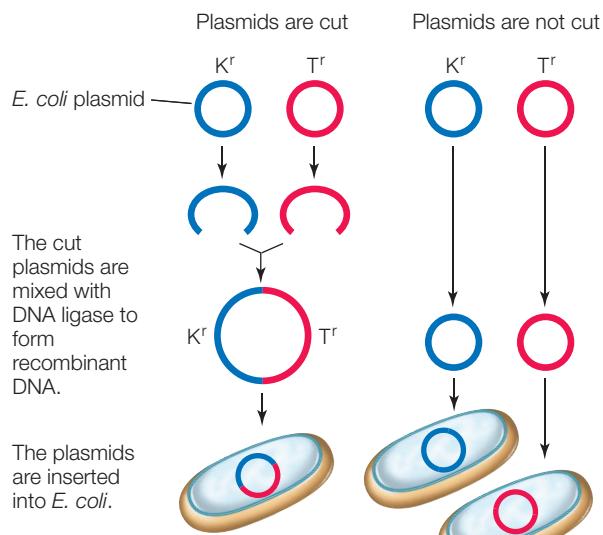
18.1 Recombinant DNA

With the discovery of restriction enzymes and DNA ligase, it became possible to combine DNA fragments from different sources in the laboratory. But would such “recombinant DNA” be functional when inserted into a living cell? The results of this experiment completely changed the scope of genetic research, increasing our knowledge of gene structure and function, and ushered in the new field of biotechnology.

HYPOTHESIS Biologically functional recombinant chromosomes can be made in the laboratory.

METHOD

E. coli plasmids carrying a gene for resistance to either the antibiotic kanamycin (K) or tetracycline (T) are cut with a restriction enzyme.



RESULTS

Some *E. coli* are resistant to both antibiotics.
No *E. coli* are doubly resistant.

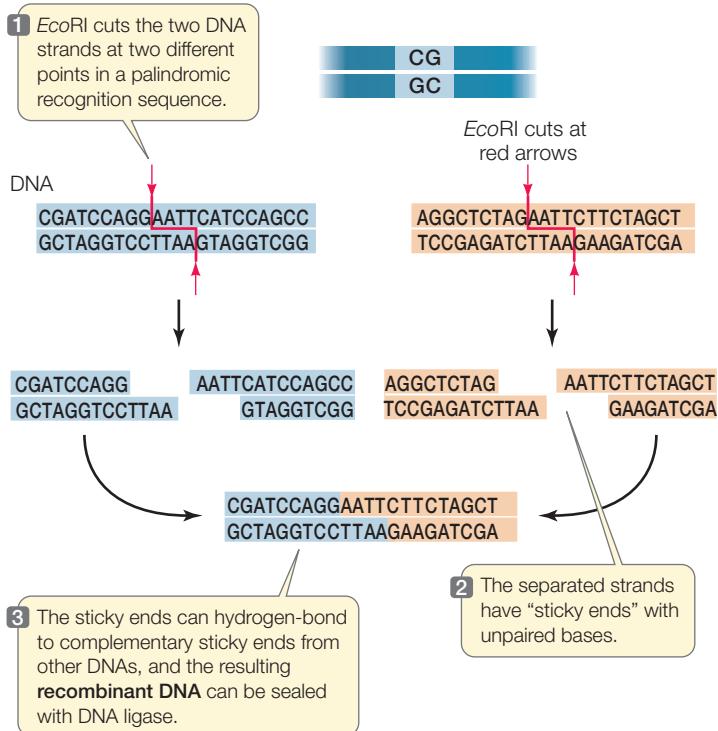
CONCLUSION

Two DNA fragments with different genes can be joined to make a recombinant DNA molecule, and the resulting DNA is functional.

FURTHER INVESTIGATION: Only one cell in 10,000 took up the plasmid in the experiment. The spontaneous mutation rate to T^r or K^r is one cell in 10⁶. How would you distinguish between genetic transformation and spontaneous mutation in this experiment?

Go to yourBioPortal.com for original citations, discussions, and relevant links for all INVESTIGATING LIFE figures.

After a DNA molecule has been cut with a restriction enzyme, complementary sticky ends can form hydrogen bonds with one another. The original ends may rejoin, or two different fragments with complementary sticky ends may join. Indeed, a fragment from one source, such as a human, can be



18.2 Cutting, Splicing, and Joining DNA Some restriction enzymes (EcoRI is shown here) make staggered cuts in DNA. EcoRI can be used to cut two different DNA molecules (blue and orange). The exposed bases can hydrogen bond with complementary exposed bases on other DNA fragments, forming recombinant DNA. DNA ligase stabilizes the recombinant molecule by forming covalent bonds in the DNA backbone.

joined to a fragment from another source, such as a bacterium. Initially the fragments are held together by weak hydrogen bonds, but then the enzyme ligase catalyzes the formation of covalent bonds between adjacent nucleotides at the ends of the fragments, joining them to form a single, larger molecule.

With these tools—restriction enzymes and DNA ligase—scientists can cut and rejoin different DNA molecules from any and all sources, including artificially synthesized DNA sequences.

18.1 RECAP

DNA fragments from different sources can be linked together to make recombinant DNA.

- How did Cohen and Boyer make the first recombinant DNA? See Figure 18.1
- How does a staggered cut in DNA create a “sticky end”? See p. 387 and Figure 18.2

Recombinant DNA has no biological significance until it is inserted inside a living cell, which can replicate and transcribe the transplanted genetic information. How can recombinant DNA made in the laboratory be inserted and expressed in living cells?

18.2 How Are New Genes Inserted into Cells?

One goal of recombinant DNA technology is to **clone**—that is, to produce many identical copies of—a particular gene. Cloning might be done for analysis, to produce a protein product in quantity, or as a step toward creating an organism with a new phenotype. Recombinant DNA is cloned by inserting it into host cells in a process known as **transformation** (or **transfection** if the host cells are derived from an animal). A host cell or organism that contains recombinant DNA is referred to as a **transgenic** cell or organism. Later in this chapter we will encounter many examples of transgenic cells and organisms, including yeast, mice, wheat plants, and even cows.

Various methods are used to create transgenic cells. Generally, these methods are inefficient in that only a few of the cells that are exposed to the recombinant DNA actually become transformed with it. In order to grow only the transgenic cells, **selectable marker** genes, such as genes that confer resistance to antibiotics, are often included as part of the recombinant DNA molecule. Antibiotic resistance genes were the markers used in Cohen and Boyer's experiment (see Figure 18.1).

Genes can be inserted into prokaryotic or eukaryotic cells

The initial successes with recombinant DNA technology were achieved using bacteria as hosts. As we have seen in preceding chapters, bacterial cells are easily grown and manipulated in the laboratory. Much of their molecular biology is known, especially for certain well-studied bacteria such as *E. coli*. Furthermore, bacteria contain plasmids, which are easily manipulated to carry recombinant DNA into the cell.

In some important ways, however, bacteria are not ideal organisms for studying and expressing eukaryotic genes. Consider how differently the processes of transcription and translation proceed in prokaryotes and eukaryotes, and recall that DNA often contains the signals for these specific functions (see Chapter 14). Furthermore, scientists often want to study how genes function in multicellular eukaryotic organisms rather than in cells grown in cultures. Or they might want to create a crop plant or farm animal with a new phenotype for use in agriculture. For these reasons, scientists have developed methods to transform or transfect eukaryotic cells.

Yeasts such as *Saccharomyces* are commonly used as eukaryotic hosts for recombinant DNA studies. The advantages of using yeasts include rapid cell division (a life cycle completed in 2–8 hours), ease of growth in the laboratory, and a relatively small genome size (about 12 million base pairs and 6,000 genes). In addition, yeasts have most of the characteristics of other eukaryotes, except for those characteristics involved in multicellularity.

Plant cells can also be used as hosts. One property that makes plant cells good hosts is the ability to make stem cells (unspecialized, totipotent cells; see Chapter 5 opener) from mature plant tissues. When these unspecialized plant cells are isolated and grown in culture, they can be transformed with recombinant DNA. These transgenic cells can be studied in culture, or manip-

ulated to form entire new plants. There are also methods for making whole transgenic plants without going through the cell culture step. These methods result in plants that carry the recombinant DNA in all their cells, including the germ line cells.

If biologists want to study expression of human or animal genes, for example for medical purposes, they use cultured animal cells as hosts. Whole transgenic animals can also be created.

Recombinant DNA enters host cells in a variety of ways

Methods for inserting DNA into host cells vary. The cells may be chemically treated to make their outer membranes more permeable, and then mixed with the DNA so that it can diffuse into the cells. Another approach is called *electroporation*; a short electric shock is used to create temporary pores in the membranes, through which the DNA can enter. Viruses can be altered so that they carry recombinant DNA into cells. Plants are often transformed using a bacterium that has evolved mechanisms to transfer its DNA into cells and then insert the DNA into a plant chromosome. Transgenic animals can be produced by injecting recombinant DNA into the nuclei of fertilized eggs. There are even "gene guns," which "shoot" the host cells with tiny particles carrying the DNA.

The challenge of inserting new DNA into a cell lies not just in getting it into the host cell, but in getting it to replicate as the host cell divides. DNA polymerase does not bind to just any sequence. If the new DNA is to be replicated, it must become part of a segment of DNA that contains an origin of replication. Such a DNA molecule is called a **replicon**, or replication unit.

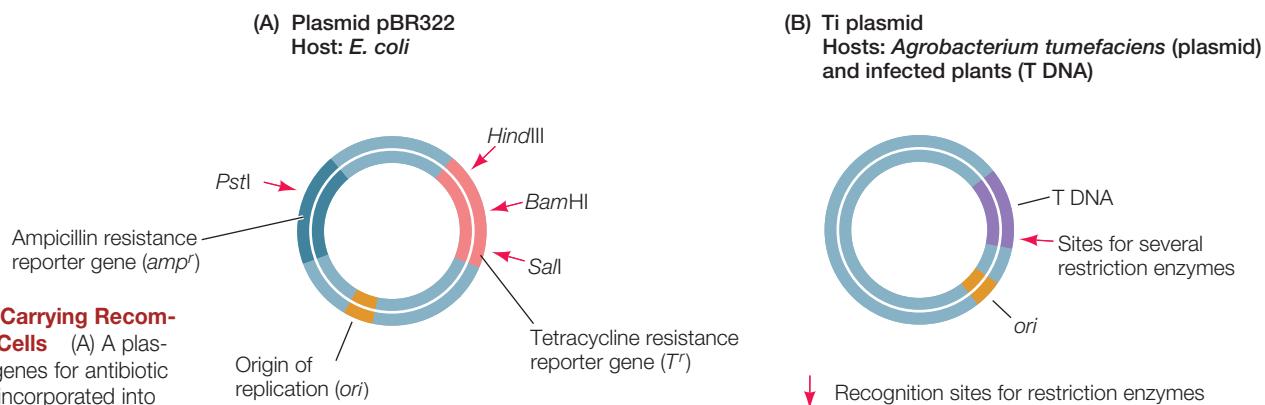
There are two general ways in which the newly introduced DNA can become part of a replicon:

- It may be inserted into a host chromosome. Although the site of insertion is usually random, this is nevertheless a common method of integrating new genes into host cells.
- It can enter the host cell as part of a carrier DNA sequence, called a **vector**, that already has an origin of replication.

Several types of vectors are used to get DNA into cells. Once inside the cells, some vectors replicate independently, while others incorporate all or part of their DNA into the host chromosomes.

PLASMIDS AS VECTORS As you learned in Chapter 12, plasmids are small chromosomes that exist in prokaryotic cells in addition to the main chromosomes. Yeast cells can also harbor plasmids. A number of characteristics make plasmids useful as transformation vectors:

- They are relatively small (an *E. coli* plasmid has 2,000–6,000 base pairs) and therefore easy to manipulate in the laboratory.
- A plasmid will usually have one or more restriction enzyme recognition sequences that each occur only once in the plasmid sequence. These sites make it easy to insert additional DNA into the plasmid before it is used to transform host cells.
- Many plasmids contain genes that confer resistance to antibiotics, which can serve as selectable markers.



- Plasmids have a bacterial origin of replication (*ori*) and can replicate independently of the host chromosome. It is not uncommon for a bacterial cell to contain hundreds of copies of a recombinant plasmid. For this reason, the power of bacterial transformation to amplify a gene is extraordinary. A one-liter culture of bacteria harboring the human β -globin gene in a typical plasmid has as many copies of that gene as the sum total of all the cells in a typical adult human being (10^{14}). A typical bacterial plasmid is shown in **Figure 18.3A**.

The plasmids used as vectors in the laboratory have been extensively altered by recombinant DNA technology to include convenient features: multiple cloning sites with 20 or more unique restriction enzyme sites for cloning purposes; origins of replication for a variety of host cells; and various kinds of reporter genes and selectable marker genes.

VIRUSES AS VECTORS Constraints on plasmid replication limit the size of the new DNA that can be inserted into a plasmid to about 10,000 base pairs. Although many prokaryotic genes may be smaller than this, most eukaryotic genes—with their introns and extensive flanking sequences—are bigger. A vector that accommodates larger DNA inserts is needed for these genes.

Both prokaryotic and eukaryotic viruses are often used as vectors for eukaryotic DNA. Bacteriophage λ , which infects *E. coli*, has a DNA genome of about 45,000 base pairs. If the genes that cause the host cell to die and lyse—about 20,000 base pairs—are eliminated, the virus can still attach to a host cell and inject its DNA. The deleted 20,000 base pairs can be replaced with DNA from another organism. Because viruses infect cells naturally, they offer a great advantage over plasmids, which often require artificial means to coax them to enter host cells. As we saw in Section 15.6, viruses are important vectors in human gene therapy.

PLASMID VECTORS FOR PLANTS An important vector for carrying new DNA into many types of plants is a plasmid found in *Agrobacterium tumefaciens*. This bacterium lives in the soil, infects plants, and causes a disease called crown gall, which is

characterized by the presence of growths, or tumors, in the plant. *A. tumefaciens* contains a plasmid called Ti (for tumor-inducing) (**Figure 18.3B**). The Ti plasmid carries genes that allow the bacterium to infect plant cells and then insert a region of its DNA called the T DNA into the chromosomes of infected cells. The T DNA contains genes that cause the growth of tumors and the production of specific sugars that the bacterium uses as sources of energy. Scientists have exploited this remarkable natural “genetic engineer” to insert foreign DNA into the genomes of plants.

When used as a vector for plant transformation, the tumor-inducing and sugar-producing genes on the T DNA are removed and replaced with foreign DNA. The altered Ti plasmids are first used to transform *Agrobacterium* cells from which the original Ti plasmids have been removed. Then the *Agrobacterium* cells are used to infect plant cells. Whole plants can be regenerated from transgenic cells or, in the case of the model plant *Arabidopsis* (see Section 17.3), the *Agrobacterium* can be used to directly infect germ line cells of whole plants.

Reporter genes identify host cells containing recombinant DNA

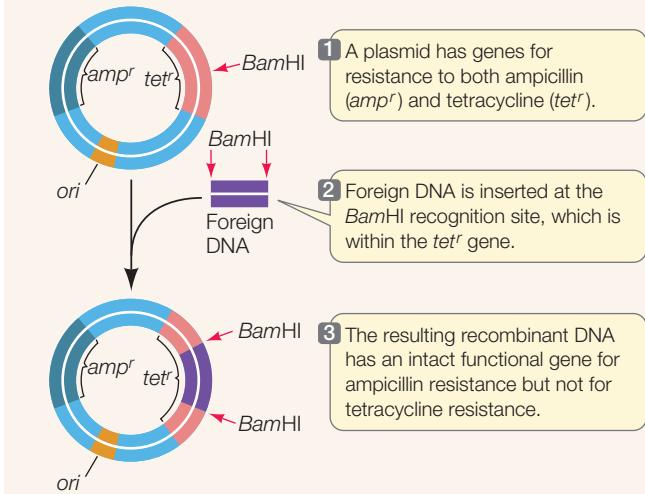
Even when a population of host cells interacts with an appropriate vector, only a small proportion of the cells actually take up the vector. Furthermore, the process of making recombinant DNA is far from perfect. After a ligation reaction, not all the vector copies contain the foreign DNA. How can we identify or select the host cells that contain that sequence?

Selectable markers such as antibiotic resistance genes can be used to select cells containing those genes. Only cells carrying the antibiotic resistance gene can grow in the presence of that antibiotic. If a vector carrying genes for resistance to two different antibiotics is used, one antibiotic can be used to selectively grow cells carrying the vector. If the other antibiotic resistance gene is inactivated by the insertion of foreign DNA, then cells carrying copies of the vector with the inserted DNA can be identified by their sensitivity to that antibiotic (**Figure 18.4**). Since the uptake of recombinant DNA is a rare event (only about 1 cell in 10,000 takes up a plasmid in such experiments), it is vital to be able to select the small number of cells harboring the recombinant DNA.

TOOLS FOR INVESTIGATING LIFE

18.4 Marking Recombinant DNA by Inactivating a Gene

Selectable marker (reporter) genes are used by scientists to select for bacteria that have taken up a plasmid. A second reporter gene allows for the identification of bacteria harboring the recombinant plasmid. The host bacteria in this experiment could display any of the three phenotypes indicated in the table.



4 Host *E. coli* are screened to detect the presence of recombinant DNA.

DNA taken up by amp^s and tet^s <i>E. coli</i> genotype	Phenotype for ampicillin	Phenotype for tetracycline
None	Sensitive	Sensitive
Foreign DNA only	Sensitive	Sensitive
pBR322 plasmid	Resistant	Resistant
pBR322 recombinant plasmid	Resistant	Sensitive

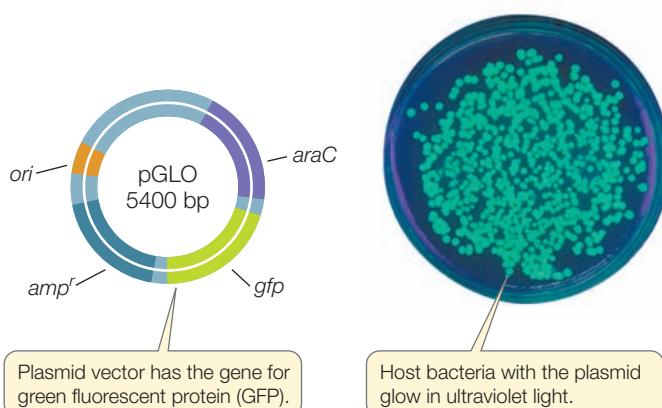
Selectable markers are one type of **reporter gene**, which is any gene whose expression is easily observed. Other reporter genes code for proteins that can be detected visually. For example:

- The β -galactosidase (lacZ) gene in the *E. coli* lac operon (see Figure 16.10) codes for an enzyme that can convert the substrate X-Gal into a bright blue product. Many plasmids contain the lacZ gene with a multiple cloning site within its sequence. Bacterial colonies containing the plasmid (which also includes an antibiotic resistance gene) are selected on a solid medium containing the antibiotic. X-Gal is also in-

cluded in the medium, so that bacterial colonies containing the recombinant DNA inserted into the lacZ gene produce white, rather than blue, colonies.

- Green fluorescent protein, which normally occurs in the jellyfish *Aequorea victoria*, emits visible light when exposed to ultraviolet light. The gene for this protein has been isolated and incorporated into vectors. It is now widely used as a reporter gene (Figure 18.5).

Such reporters are not just used to select and identify cells carrying recombinant DNA. They can be attached to promoters in order to study how the promoters function under different conditions or in different tissues of a transgenic multicellular organism. They can also be attached to other proteins, to study how and where those proteins become localized within eukaryotic cells.



- 18.5 Green Fluorescent Protein as a Reporter** The presence of a plasmid with the gene for green fluorescent protein is readily apparent in transgenic cells because they glow under ultraviolet light. This allows the identification of cells carrying a plasmid without the use of selection on antibiotics. That is, no cells are killed during the selection process.

18.2 RECAP

Recombinant DNA can be cloned by using a vector to insert it into a suitable host cell. The vector often has a selectable marker or other reporter gene that gives the host cell a phenotype by which transgenic cells can be identified.

- List the characteristics of a plasmid that make it suitable for introducing new DNA into a host cell. See pp. 389–390
- How are cells harboring a vector that carries recombinant DNA selected? See p. 390 and Figure 18.4

We have described how DNA can be cut or amplified, inserted into a vector, and introduced into host cells. We have also seen how host cells carrying recombinant DNA can be identified. Now let's consider where the genes or DNA fragments used in these procedures come from.

18.3 What Sources of DNA Are Used in Cloning?

A major goal of cloning experiments is to elucidate the functions of DNA sequences and the proteins they encode. The DNA fragments used in cloning procedures are obtained from a number of sources. They include random fragments of chromosomes that are maintained as gene libraries, complementary DNA obtained by reverse transcription from mRNA, products of the polymerase chain reaction (PCR), and artificially synthesized or mutated DNA.

Often a scientist will want to express a gene derived from one kind of organism in another, very different organism—for example, a human gene in a bacterium, or a bacterial gene in a plant. To do this it is necessary to use a promoter and other regulatory sequences from the host organism: a bacterial promoter will not function in a plant cell, for example. The coding region of the gene of interest is inserted between a promoter and a transcription termination sequence derived from the host organism, or from one that uses similar mechanisms for gene regulation.

Libraries provide collections of DNA fragments

In Chapter 17 we introduced the concept of a **genomic library**: a collection of DNA fragments that together comprise the genome of an organism. Now we provide details on how a genomic or other gene library is generated and used.

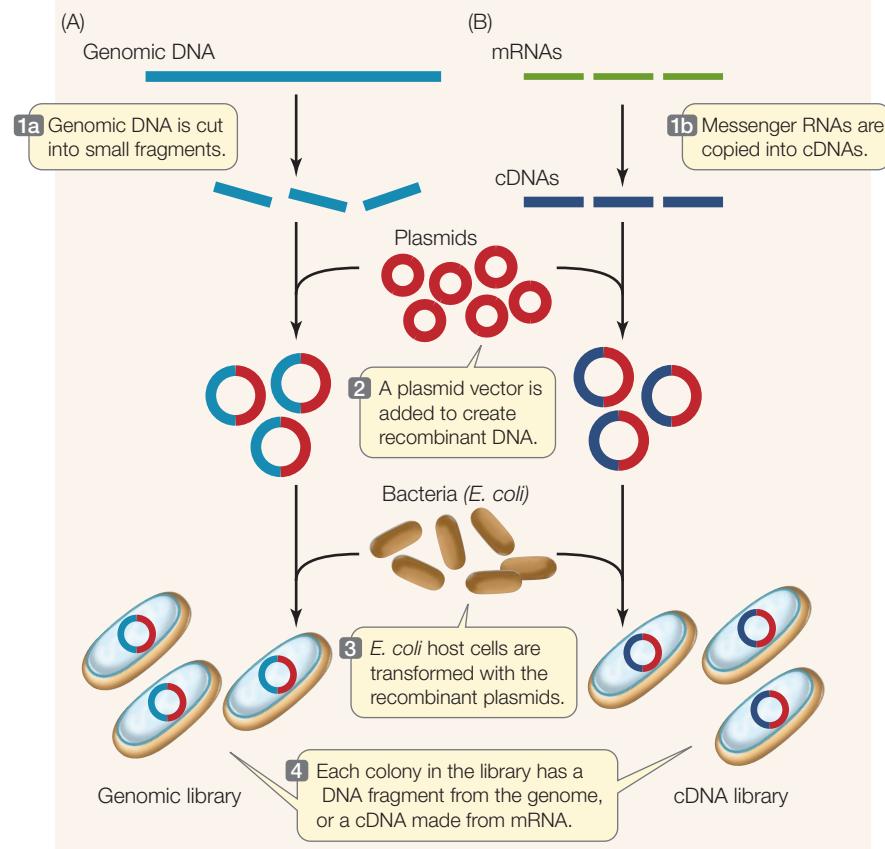
Restriction enzymes or other means, such as mechanical shearing, can be used to break chromosomes into smaller pieces. These smaller DNA fragments still constitute a genome (**Figure 18.6A**), but the information is now in many smaller “volumes.” Each fragment is inserted into a vector, which is then taken up by a host cell. Proliferation of a single transformed cell produces a colony of recombinant cells, each of which harbors many copies of the same fragment of DNA.

When plasmids are used as vectors, about 700,000 separate fragments are required to make a library of the human genome. By using bacteriophage λ , which can carry four times as much DNA as a plasmid, the number of “volumes” in the library can be reduced to about 160,000. Although this seems like a large number, a single petri plate can hold thousands of phage colonies, or plaques, and is easily screened for the presence of a particular DNA sequence by hybridization to an appropriate DNA probe.

TOOLS FOR INVESTIGATING LIFE

18.6 Constructing Libraries

Intact genomic DNA is too large to be introduced into host cells. A genomic library can be made by breaking the DNA into small fragments, incorporating the fragments into a vector, and then transforming host cells with the recombinant vectors. Each colony of cells contains many copies of a small part of the genome. Similarly, there are many mRNAs in a cell. These can be copied into cDNAs and a library made from them. The DNA in these colonies can then be isolated for analysis.



cDNA libraries are constructed from mRNA transcripts

A much smaller DNA library—one that includes only the genes transcribed in a particular tissue—can be made from **complementary DNA**, or **cDNA** (**Figure 18.6B**). This involves isolating mRNA from cells, then making cDNA copies of that mRNA by complementary base pairing. An enzyme, reverse transcriptase, catalyzes this reaction.

A collection of cDNAs from a particular tissue at a particular time in the life cycle of an organism is called a **cDNA library**. Messenger RNAs do not last long in the cytoplasm and are often present in small amounts, so a cDNA library is a “snapshot” that preserves the transcription pattern of the cell. Complementary DNA libraries have been invaluable for comparing gene expression in different tissues at different stages of development. For example, their use has shown that up to one-third of all the genes of an animal are expressed only during development. Complementary DNA is also a good starting point for cloning eukaryotic genes (because the clones contain only the

coding sequences of the genes) and genes that are expressed in only a few cell types.

Synthetic DNA can be made by PCR or by organic chemistry

In Chapter 13 we describe the polymerase chain reaction (PCR), a method of amplifying DNA in a test tube. PCR can begin with as little as 10^{-12} g of DNA (a picogram). Any fragment of DNA can be amplified by PCR as long as appropriate primers are available. You will recall that DNA replication (by PCR or any other system) requires not just a template on which DNA polymerase adds complementary nucleotides, but also a short oligonucleotide primer where replication begins (see Figure 13.22). If the appropriate primers (two are needed—one for each strand of DNA) are added to denatured DNA, more than two billion copies of the DNA region between the primers can be produced in just a few hours. This amplified DNA can then be inserted into plasmids to create recombinant DNA and cloned in host cells.

The *artificial synthesis* of DNA by organic chemistry is now fully automated, and a special service laboratory can make short- to medium-length sequences overnight for any number of investigators. Synthetic oligonucleotides (single-stranded DNA fragments of up to 40 bp) are used as primers in PCR reactions. These primers can be designed to create short new sequences at the ends of the PCR products. This might be done to create a mutation in a recombinant gene, or to add restriction enzyme sites at the ends of the PCR product to aid in cloning.

Longer synthetic sequences can be pieced together to construct an artificial gene. If we know the amino acid sequence of the desired protein product, we can use the genetic code to figure out the corresponding DNA sequence. As mentioned above, other sequences must be added, such as the promoter and transcription termination sequences. Appropriate selection of the codon for a given amino acid is another important consideration: many amino acids are encoded by more than one codon (see Figure 14.6), and host organisms vary in their use of synonymous codons.

DNA mutations can be created in the laboratory

Mutations that occur in nature have been important in demonstrating cause-and-effect relationships in biology. However, mutations in nature are rare events. Recombinant DNA technology allows us to ask “what if” questions by creating mutations artificially. Because synthetic DNA can be made with any desired sequence, it can be manipulated to create specific mutations, the consequences of which can be observed when the mutant DNA is expressed in host cells. These mutagenesis techniques have revealed many cause-and-effect relationships.

For example, consider the experiment illustrated in Figure 14.20. Researchers hypothesized that a nuclear localization signal (NLS) sequence of amino acids is necessary for targeting a protein to the nucleus after it is made at the ribosome. The researchers used recombinant DNA technology to synthesize genes encoding proteins with and without the sequence, which

were then used to transform cells. Without the NLS, newly synthesized proteins did not enter the nucleus. Knowing this, the researchers then asked, “Are certain amino acids more functionally important to the NLS than others?” In follow-up experiments, they made a series of mutated genes to test whether certain amino acids were needed at certain locations in the NLS. They found that changing the amino acids at the very beginning or very end of the NLS, but not the middle, abolished its function. This led to a fuller description of the binding of the NLS to its nuclear receptor. Without the ability to generate specific mutations, these experiments would not have been possible.

18.3 RECAP

DNA for cloning can be obtained from genomic libraries, cDNA made from mRNA, or artificially synthesized DNA fragments. Gene function can be investigated by intentionally introducing mutations into natural or synthetic genes and organisms.

- How are genomic DNA and cDNA libraries made and used? **See p. 392 and Figure 18.6**
- Explain how recombinant DNA and mutagenesis are used to test “what if” questions in biology. **See p. 393**

We’ve explored the various sources of DNA that can be used to make recombinant DNA molecules and the ways the resulting molecules can be used to study the functions of genes and proteins. We now turn to some additional tools that are available for studying DNA.

18.4 What Other Tools Are Used to Study DNA Function?

Sections 13.5 and 17.1 describe PCR and DNA sequencing, two important techniques arising from our understanding of DNA replication. In this section we will examine three additional techniques for studying DNA, including homologous recombination to inactivate genes, antisense and RNAi to block gene expression, and DNA microarrays to analyze large numbers of nucleotide sequences.

Genes can be inactivated by homologous recombination

One way to study a gene or protein in order to understand its function is to inactivate the gene so that it is not transcribed and translated into a functional protein. Such a manipulation is called a **knockout** experiment. In plants, transposons or T DNA insertions can be used to create thousands of knockout mutants, and then the mutants are screened to identify those with altered phenotypes. For example, the mutants can be screened for those that are susceptible to a particular disease. This is an important way to identify genes that are involved in processes such as resistance to disease or other environmental stresses, such as drought and temperature extremes.

A technique called **homologous recombination** is a much more targeted way to produce knockout mutants. In this case, the gene of interest has already been identified, and recombinant DNA technology is used to specifically inactivate that gene. Mice are frequently used in such knockout experiments (**Figure 18.7**). The normal allele of the mouse gene to be tested is inserted into a plasmid. Restriction enzymes are then used to insert a fragment containing a reporter gene into the middle of the normal gene. This addition of extra DNA plays havoc with the targeted gene's transcription and translation; a functional mRNA is seldom made from a gene whose sequence has been thus interrupted.

Once the recombinant plasmid has been made, it is used to transfect mouse embryonic stem cells. (A **stem cell** is an unspecialized cell that divides and differentiates into specialized cells.) Much of the targeted gene is still present in the plasmid (although in two separated regions), and these sequences tend to line up with their homologous sequences in the normal allele on the mouse chromosome. Sometimes recombination occurs, and the plasmid's inactive allele is "swapped" with the functional allele in the host cell. The inactive allele is inserted permanently into the host cell's genome and the normal allele is lost (because the plasmid cannot replicate in mouse cells). The active reporter gene in the insert is used to select those stem cells carrying the inactivated gene.

A transfected stem cell is now transplanted into an early mouse embryo. If the mouse that develops from this embryo has the mutant gene in its germ line cells, its progeny will have the knockout gene in every cell in their bodies. Such mice are inbred to create *knockout mice* carrying the inactivated gene in homozygous form. The mutant mouse can then be observed for phenotypic changes, to find clues about the function of the targeted gene in the normal (wild-type) animal. The knockout technique has been important in assessing the roles of many genes, and has been especially valuable in studying human genetic diseases. Many of these diseases, such as phenylketonuria, have knockout mouse *models*—mouse strains that suffer from an analogous disease—produced by homologous recombination. These models can be used to study a disease and to test potential treatments. Mario Capecchi, Martin Evans, and Oliver Smithies shared the Nobel Prize for developing the knockout mouse technique.

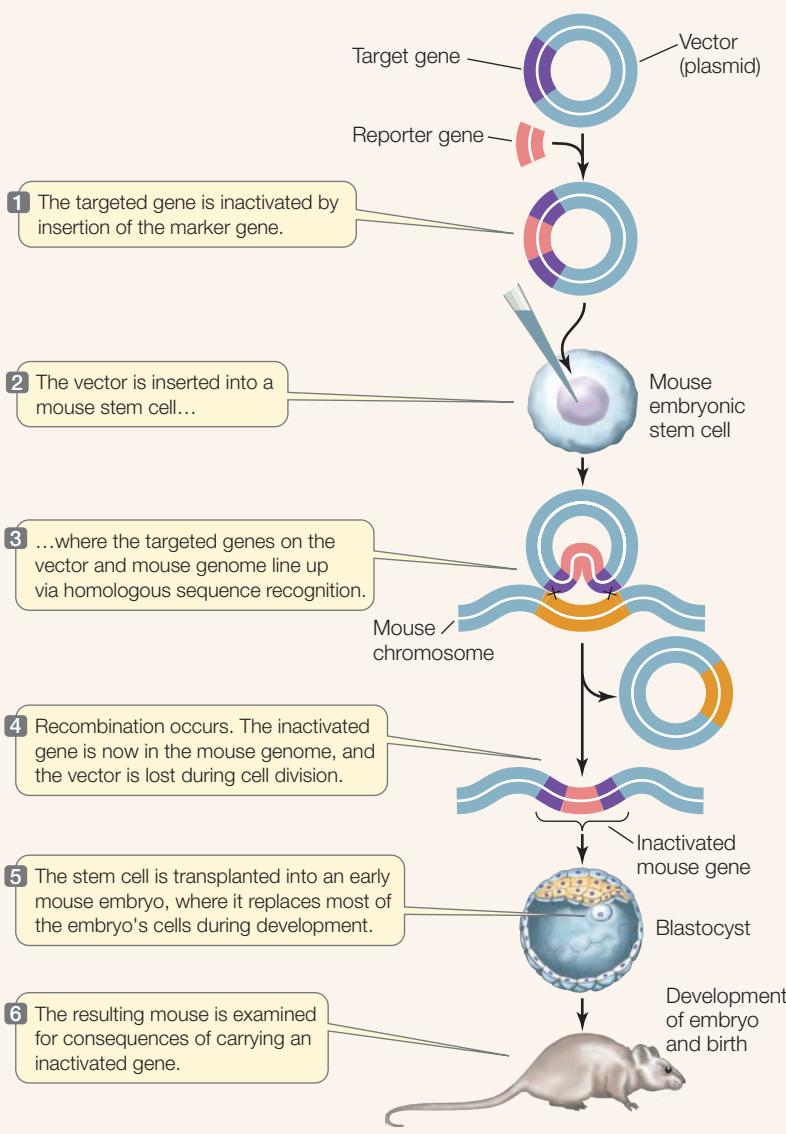
Complementary RNA can prevent the expression of specific genes

Another way to study the expression of a specific gene is to block the translation of its mRNA. This is an example of scientists imitating nature. As described in Section 16.5, gene expres-

TOOLS FOR INVESTIGATING LIFE

18.7 Making a Knockout Mouse

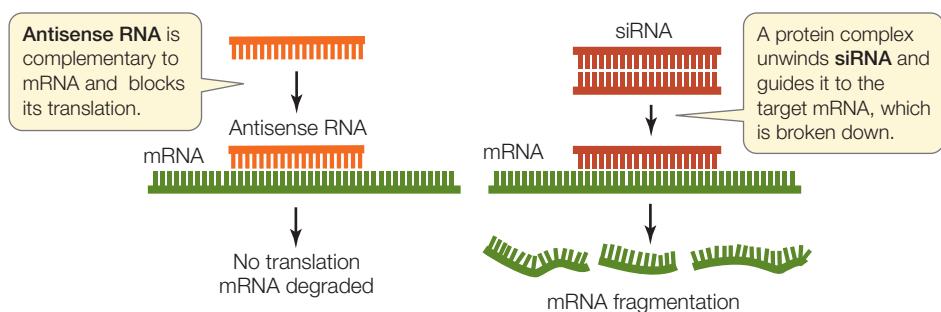
Animals carrying mutations are rare. Homologous recombination is used to replace a normal mouse gene with an inactivated copy of that gene, thus "knocking out" the gene. Discovering what happens to a mouse with an inactive gene tells us much about the normal role of that gene.



sion is sometimes controlled by the production of double-stranded RNA molecules, which are cut up and unwound to produce short, single-stranded RNA molecules (microRNAs) that are complementary to specific mRNA sequences (see Figure 16.23). Such a complementary molecule is called **antisense RNA** because it binds by base pairing to the "sense" bases on the mRNA. The resulting partially double-stranded RNA hybrid inhibits translation of the mRNA, and the hybrid tends to be broken down rapidly in the cytoplasm. Although the gene continues to be transcribed, translation does not take place. After determining the sequence of a gene and its mRNA in the lab-

18.8 Using Antisense RNA and siRNA to Block Translation of mRNA

Once a gene's sequence is known, the synthesis of its protein can be prevented by making either an antisense RNA (left) or a small interfering RNA (siRNA, right) that is complementary to its mRNA.



oratory, scientists can make a specific, single-stranded antisense RNA and add it to a cell to prevent translation of that gene's mRNA (**Figure 18.8, left**).

Several antisense drugs are being developed to reduce the expression of genes involved with cancer. For example, the gene *bcl2* codes for a protein that blocks apoptosis, and in some forms of cancer *bcl2* is activated inappropriately through mutation. These cells fail to undergo apoptosis, continue to divide, and form a tumor. Treatment with oblimersen, an antisense RNA that binds to *bcl2* mRNA, prevents production of the protein, and leads to apoptosis of tumor cells and shrinkage of the tumor.

A related technique to antisense RNA takes advantage of **RNA interference (RNAi)**, a rare, natural mechanism for inhibiting mRNA translation. In a process similar to that involved in processing microRNAs, a short (about 20 nucleotides) double-stranded RNA is unwound to single strands by a protein complex that guides this RNA to a complementary region on mRNA. The protein complex catalyzes the breakdown of the targeted mRNA. RNAi was not discovered until the late 1990s, but since then scientists have synthesized double-stranded siRNAs to inhibit the expression of known genes (**Figure 18.8, right**). Because these double-stranded siRNAs are more stable than antisense RNAs, the use of siRNAs is the preferred approach for blocking translation. Macular degeneration is an eye disease that results in near-blindness when blood vessels proliferate in the eye. The signaling molecule that stimulates vessel proliferation is a growth factor. An RNAi-based therapy is being developed to target this growth factor's mRNA and shows promise in stopping and even reversing the progress of the disease.

Although medical applications for RNAi are still at the experimental stage, antisense RNA and RNAi have been widely used to test cause-and-effect relationships in biological research. Another powerful research tool with great potential for medicine is the gene chip, or DNA microarray.

DNA microarrays can reveal RNA expression patterns

The emerging science of genomics has to face two major quantitative realities. First, there are very large numbers of genes in eukaryotic genomes. Second, the pattern of gene expression in different tissues at different times is quite distinctive. For example, a cell from a skin cancer at its early stage may have a unique mRNA "fingerprint" that differs from that of both normal skin cells and the cells of a more advanced skin cancer.

To find such patterns, scientists could isolate mRNA from a cell and test it by hybridization with each gene in the genome, one gene at a time. But that would involve many steps and take a very long time. It is far simpler to do these hybridizations all in one step. This is possible with **DNA microarray** technology, which provides large arrays of sequences for hybridization experiments.

The development of DNA arrays ("gene chips") was inspired by methods used for decades by the semiconductor industry. A silicon microchip consists of an array of microscopic electric circuits etched onto a tiny silicon base, called a chip. In the same way, a series of DNA sequences can be attached to a glass slide in a precise order (**Figure 18.9**). The slide is divided into a grid of microscopic spots, or "wells." Each spot contains thousands of copies of a particular oligonucleotide of 20 or more bases. A computer controls the addition of these oligonucleotide sequences in a predetermined pattern. Each oligonucleotide can hybridize with only one DNA or RNA sequence, and thus is a unique identifier of a gene. Many thousands of different oligonucleotides can be placed on a single microarray.

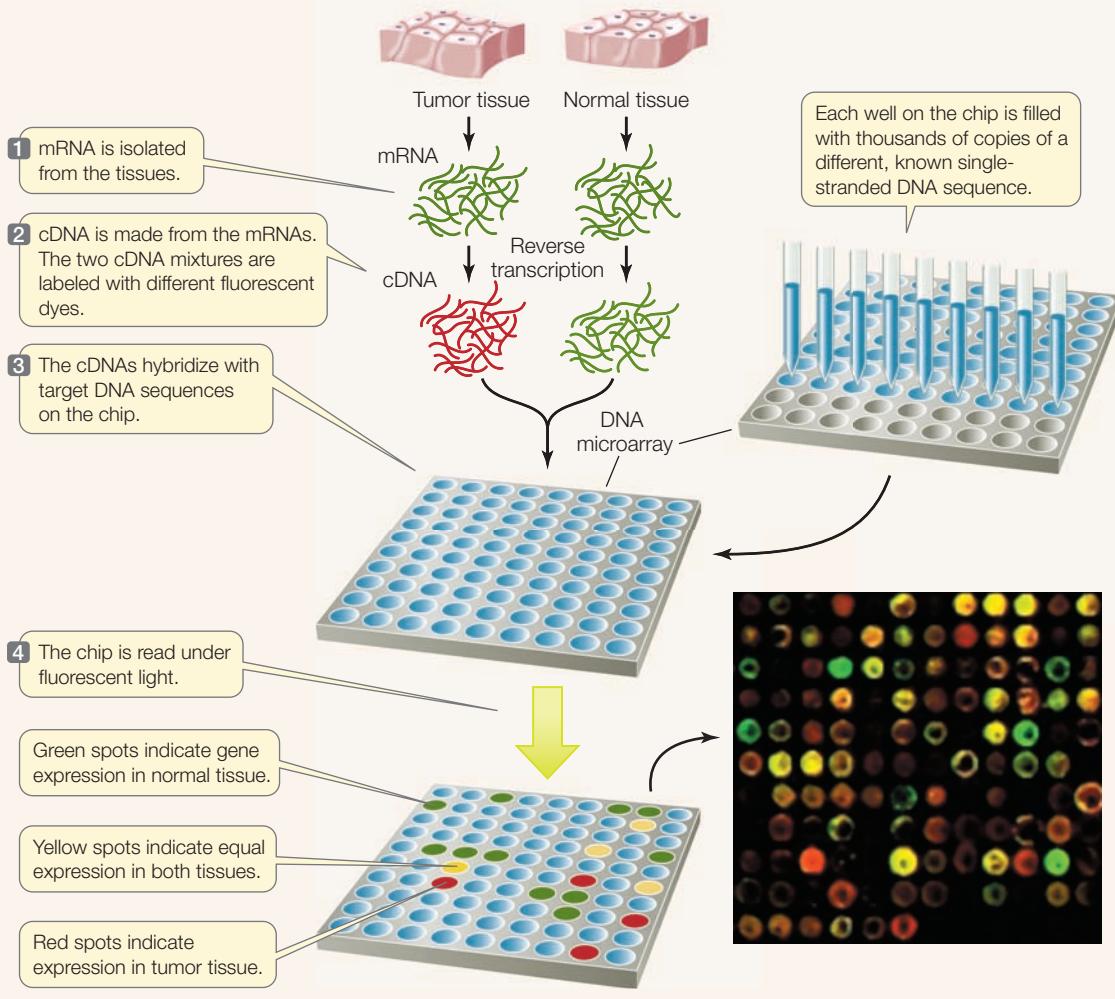
As we mention in Section 17.4, DNA microarrays can be used to identify specific single nucleotide polymorphisms or other mutations in genomic DNA samples. Or they can be used to analyze RNA from different tissues or cells to identify which genes are expressed in those cell types. If mRNA is to be analyzed, it is usually incubated with reverse transcriptase to make cDNA (see Figure 18.6B). Fluorescent dyes are used to tag the cDNAs from different samples with different colors (usually red and green; see Figure 18.9). The cDNAs are used to probe the DNA on the microarray. Complementary sequences that form hybrids with the DNA on the microarray can be located using a sensitive scanner that detects the fluorescent light.

A clinical use of DNA chips was developed by Laura van 't Veer and her colleagues at the Netherlands Cancer Institute (**Figure 18.10**). Most women with breast cancer are treated with surgery to remove the tumor, and then treated with radiation soon afterward to kill cancer cells that the surgery may have missed. But a few cancer cells may still survive in some patients, and these eventually form tumors in the breast or elsewhere in the body. The challenge for physicians is to develop criteria to identify patients with surviving cancer cells so that they can be treated aggressively with tumor-killing chemotherapy. The scientists in van 't Veer's group followed the medical histories of breast cancer patients to identify those patients whose cancer recurred. They then used a DNA microarray to examine the expression of about 1,000 genes in these patients' original tu-

TOOLS FOR INVESTIGATING LIFE

18.9 DNA on a Chip

Large arrays of DNA sequences can be used to identify specific sequences in a sample of DNA or RNA by hybridization. For example, thousands of known, synthetic DNA sequences can be attached to a glass slide in an organized grid pattern. This can be hybridized with cDNA samples derived from two different tissues to find out what genes are being expressed in the tissues.

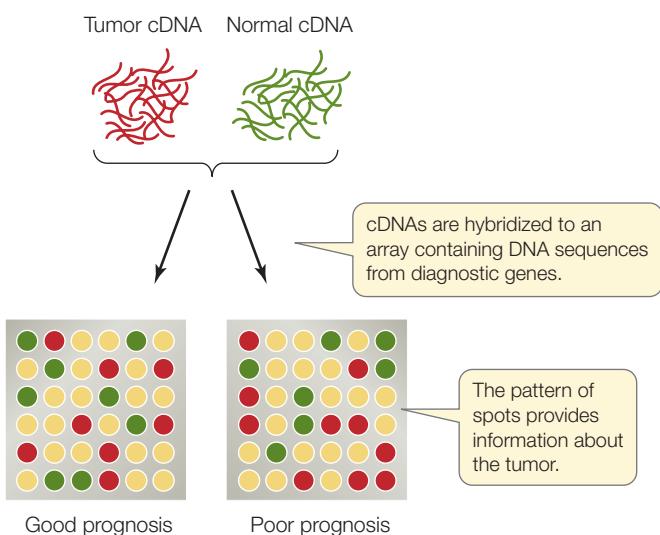


yourBioPortal.com

GO TO Animated Tutorial 18.1 • DNA Chip Technology

mors (which had been stored after their surgical removal) relative to normal tissue. They found 70 genes whose expression differed dramatically between tumors from patients whose cancers recurred and tumors from patients whose cancers did not recur. From this information, the Dutch group was able to identify what is called a *gene expression signature*. This expression pattern is useful in clinical decision-making: patients with a good prognosis can avoid unnecessary chemotherapy, while those with a poor prognosis can receive aggressive treatment.

18.10 Using DNA Arrays for Medical Diagnosis The pattern of expression of 70 genes in tumor tissues indicates whether breast cancer is likely to recur. Actual arrays have more dots than shown here.



18.4 RECAP

Researchers can study the function of a gene by knocking out that gene in a living organism. Anti-sense RNAs and siRNAs silence genes by selectively blocking mRNA translation. DNA microarrays allow the simultaneous analysis of many different mRNA transcripts.

- How is a gene “knocked out” in a living organism? See p. 393 and Figure 18.7
- How do antisense RNA and siRNA molecules affect gene expression? See pp. 394–395 and Figure 18.9

Now that you’ve seen how DNA can be fragmented, recombined, manipulated, and put back into living organisms, let’s see some examples of how these techniques are used to make useful products.

18.5 What Is Biotechnology?

Biotechnology is the use of cells or whole living organisms to produce materials useful to people, such as foods, medicines, and chemicals. People have been doing this for a very long time. For example, the use of yeasts to brew beer and wine dates back at least 8,000 years, and the use of bacterial cultures to make cheese and yogurt is a technique many centuries old. For a long time people were not aware of the molecular basis of each of these biochemical transformations.

About 100 years ago, thanks largely to Louis Pasteur’s work, it became clear that specific bacteria, yeasts, and other microbes could be used as biological converters to make certain products. Alexander Fleming’s discovery that the mold *Penicillium* makes the antibiotic penicillin led to the large-scale commercial culture of microbes to produce antibiotics as well as other useful chemicals. Today, microbes are grown in vast quantities to make much of the industrial-grade alcohol, glycerol, butyric acid, and citric acid that are used by themselves or as starting materials in the manufacture of other products.

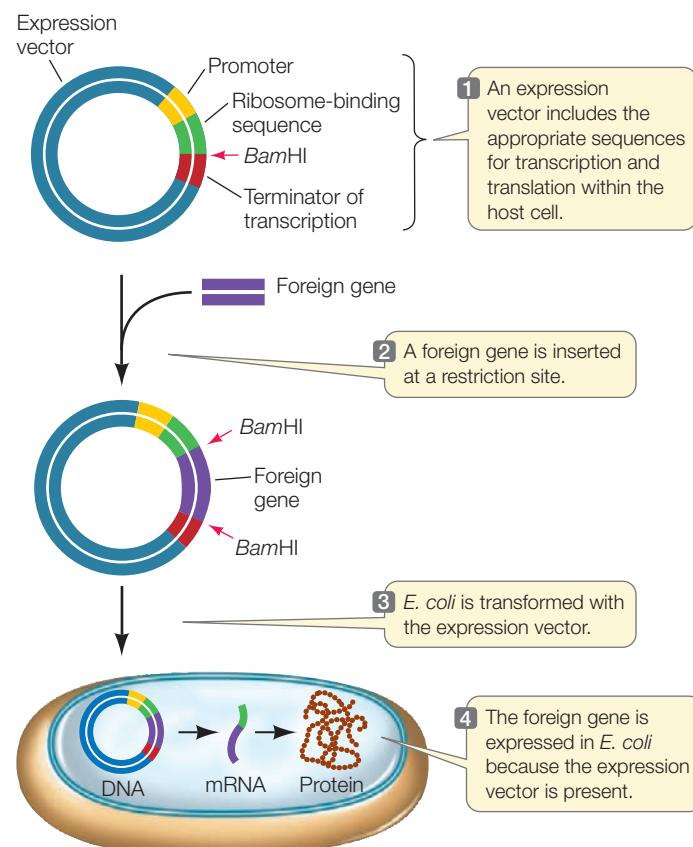
Nevertheless the commercial harvesting of proteins, including hormones and enzymes, was limited by the (often) minuscule amounts that could be extracted from organisms that produce them naturally. Yields were low, and purification was difficult and costly. Gene cloning has changed all this. The abil-

ity to insert almost any gene into bacteria or yeasts, along with methods to induce the gene to make its product in large amounts and export it from the cells, has turned these microbes into versatile factories for important products. Today there is interest in producing nutritional supplements and pharmaceuticals in whole transgenic animals and harvesting them in large quantities, for example from the milk of cows or the eggs of chickens. Key to this boom in biotechnology has been the development of specialized vectors that not only carry genes into cells, but also make those cells express them at high levels.

Expression vectors can turn cells into protein factories

If a eukaryotic gene is inserted into a typical plasmid and used to transform *E. coli*, little if any of the gene product will be made. Other key prokaryotic DNA sequences must be included with the gene. A bacterial promoter, a signal for transcription termination, and a special sequence that is necessary for ribosome binding on the mRNA must all be included in the transformation vector if the gene is to be expressed in the bacterial cell.

To solve this kind of problem, scientists make **expression vectors** that have all the characteristics of typical vectors, as well as the extra sequences needed for the foreign gene (also called a *transgene*) to be expressed in the host cell. For bacterial hosts, these additional sequences include the elements named above (Figure 18.11); for eukaryotes, they include the poly A-addition sequence, transcription factor binding sites, and enhancers. An expression vector can be designed to deliver transgenes to



18.11 Expression of a Transgene in a Host Cell Produces Large Amounts of its Protein Product

To be expressed in *E. coli*, a gene derived from a eukaryote requires bacterial sequences for transcription initiation (promoter), transcription termination, and ribosome binding. Expression vectors contain these additional sequences, enabling the eukaryotic protein to be synthesized in the prokaryotic cell.

any class of prokaryotic or eukaryotic host and may include additional features:

- An *inducible promoter*, which responds to a specific signal, can be included. For example, a promoter that responds to hormonal stimulation can be used so that the transgene will be expressed at high levels when the hormone is added.
- A *tissue-specific promoter*, which is expressed only in a certain tissue at a certain time, can be used if localized expression is desired. For example, many seed proteins are expressed only in the plant embryo. Coupling a transgene to a seed-specific promoter will allow it to be expressed only in seeds.
- *Signal sequences* can be added so that the gene product is directed to an appropriate destination. For example, when a protein is made by yeast or bacterial cells in a liquid medium, it is economical to include a signal directing the protein to be secreted into the extracellular medium for easier recovery.

18.5 RECAP

Expression vectors maximize the expression of transgenes inserted into host cells.

- How do expression vectors work? See pp. 397–398 and Figure 18.11

This chapter has introduced many of the methods that are used in biotechnology. Let's turn now to the ways biotechnology is being applied to meet some specific human needs.

18.6 How Is Biotechnology Changing Medicine, Agriculture, and the Environment?

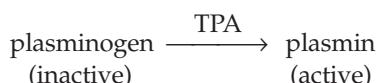
Huge potential for improvements in health, agriculture, and the environment derive from recent developments in biotechnology. We now have the ability to make virtually any protein by recombinant DNA technology and to insert transgenes into many kinds of host cells. With these revolutionary develop-

ments in biological capability, concerns have been raised about ethics and safety. We now turn to the promises and problems of biotechnology that uses DNA manipulation.

Medically useful proteins can be made by biotechnology

Many medically useful products are being made by biotechnology (Table 18.1), and hundreds more are in various stages of development. The manufacture of *tissue plasminogen activator* (TPA) provides a good illustration of a medical application of biotechnology.

When a wound begins bleeding, a blood clot soon forms to stop the flow. Later, as the wound heals, the clot dissolves. How does the blood perform these conflicting functions at the right times? Mammalian blood contains an enzyme called *plasminogen*. When activated, it becomes *plasmin* and catalyzes the dissolution of the clotting proteins. The conversion of plasminogen to plasmin is catalyzed by the enzyme TPA, which is produced by cells lining the blood vessels:



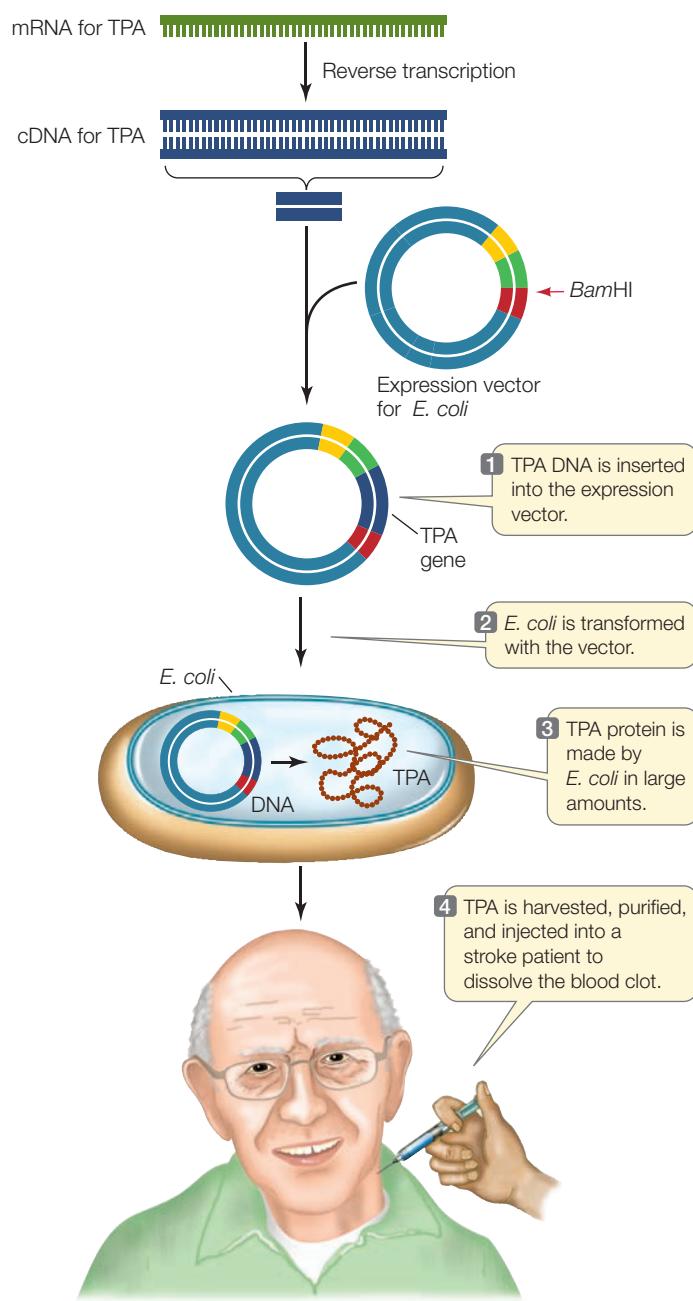
Heart attacks and strokes can be caused by blood clots that form in major blood vessels leading to the heart or the brain, respectively. During the 1970s, a bacterial enzyme called *streptokinase* was found to stimulate the dissolution of clots in some patients. Treatment with this enzyme saved lives, but being a foreign protein, it triggered the body's immune system to react against it. More important, the drug sometimes prevented clotting throughout the entire circulatory system, sometimes leading to a dangerous situation in which blood could not clot where needed.

When TPA was discovered, it had many advantages: it bound specifically to clots, and it did not provoke an immune reaction. But the amounts of TPA that could be harvested from human tissues were tiny, certainly not enough to inject at the site of a clot in the emergency room.

Recombinant DNA technology solved this problem. TPA mRNA was isolated and used to make cDNA, which was then

TABLE 18.1
Some Medically Useful Products of Biotechnology

PRODUCT	USE
Colony-stimulating factor	Stimulates production of white blood cells in patients with cancer and AIDS
Erythropoietin	Prevents anemia in patients undergoing kidney dialysis and cancer therapy
Factor VIII	Replaces clotting factor missing in patients with hemophilia A
Growth hormone	Replaces missing hormone in people of short stature
Insulin	Stimulates glucose uptake from blood in people with insulin-dependent (Type I) diabetes
Platelet-derived growth factor	Stimulates wound healing
Tissue plasminogen activator	Dissolves blood clots after heart attacks and strokes
Vaccine proteins: Hepatitis B, herpes, influenza, Lyme disease, meningitis, pertussis, etc.	Prevent and treat infectious diseases



18.12 Tissue Plasminogen Activator: From Protein to Gene to Drug

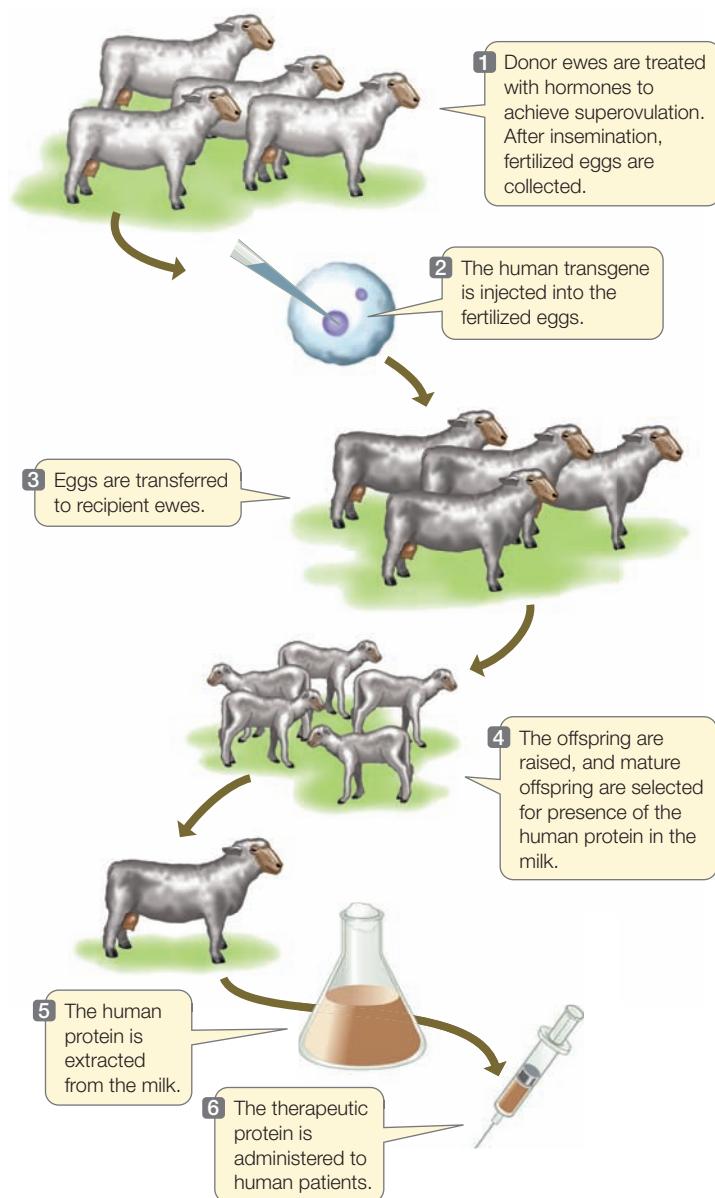
TPA is a naturally occurring human protein that dissolves blood clots. It is used to treat patients suffering from blood clotting in heart attacks or strokes, and is manufactured using recombinant DNA technology.

inserted into an expression vector and used to transform *E. coli* (Figure 18.12). The transgenic bacteria made the protein in quantity, and it soon became available commercially. This drug has had considerable success in dissolving blood clots in people experiencing strokes and heart attacks.

Another way of making medically useful products in large amounts is **pharming**: the production of pharmaceuticals in farm animals or plants. For example, a gene encoding a useful protein might be placed next to the promoter of the gene that encodes lactoglobulin, an abundant milk protein. Transgenic animals car-

rying this recombinant DNA will secrete large amounts of the foreign protein into their milk. These natural “bioreactors” can produce abundant supplies of the protein, which can be separated easily from the other components of the milk (Figure 18.13).

Human growth hormone is a protein made in the pituitary gland in the brain and has many effects, especially in growing children (see Chapter 41). Children with growth hormone deficiency have short stature as well as other abnormalities. In the past they were treated with protein isolated from the pituitary glands of dead people, but the supply was too limited to meet demand. Recombinant DNA technology was used to coax bacteria to make the protein, but the cost of treatment was high (\$30,000 a year). In 2004, a team led by Daniel Salamone at the



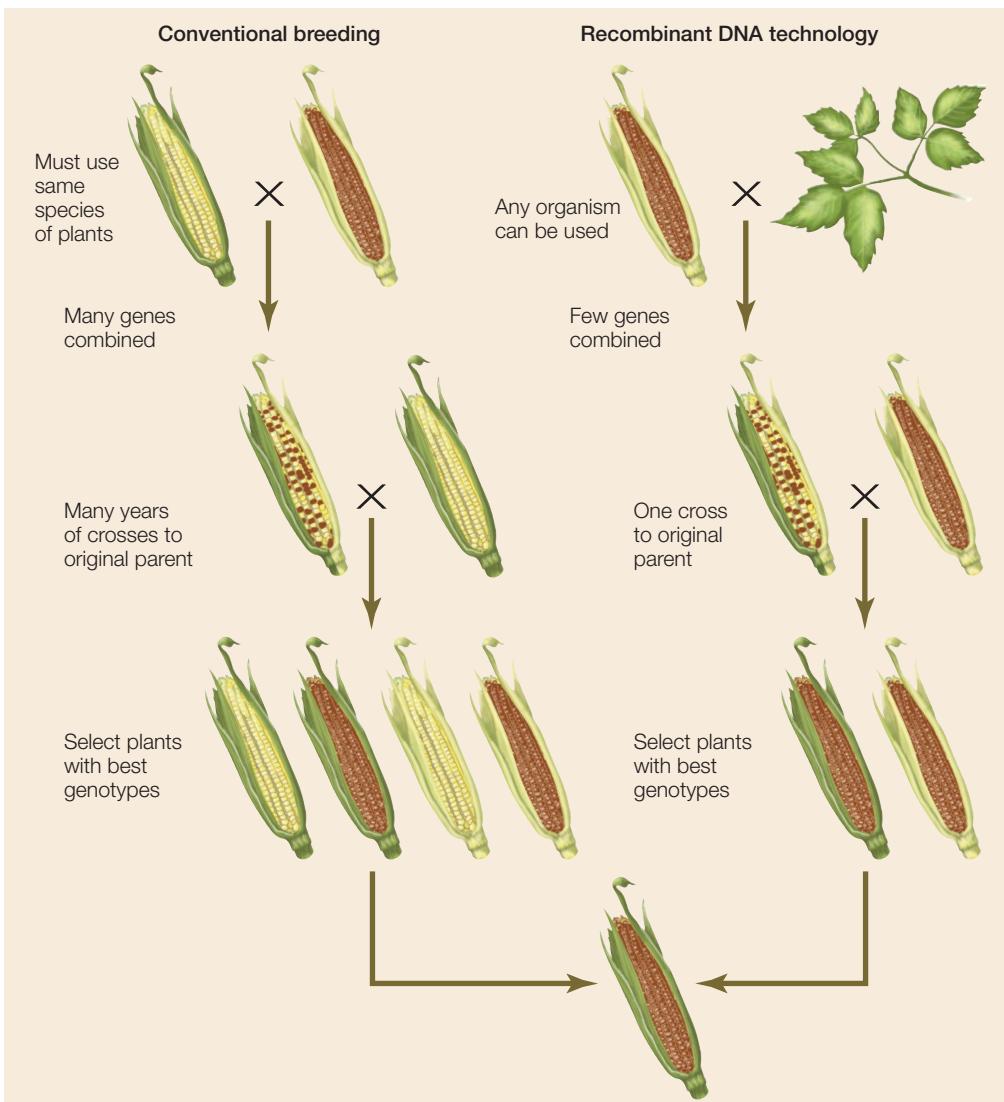
18.13 Pharming An expression vector carrying a desired gene can be put into an animal egg, which is implanted into a surrogate mother. The transgenic offspring produce the new protein in their milk. The milk is easily harvested and the protein isolated, purified, and made clinically available to patients.

University of Buenos Aires made a transgenic cow that secretes human growth hormone in her milk. The yield is prodigious: only 15 such cows are needed to meet the needs worldwide of children suffering from this type of dwarfism.

DNA manipulation is changing agriculture

The cultivation of plants and the husbanding of animals provide the world's oldest examples of biotechnology, dating back more than 10,000 years. Over the centuries, people have adapted crops and farm animals to their needs. Through selective breeding of these organisms, desirable characteristics such as large seeds, high fat content in milk, or resistance to disease have been imparted and improved.

Until recently, the most common way to improve crop plants and farm animals was to identify individuals with desirable phenotypes that existed as a result of natural variation. Through many deliberate crosses, the genes responsible for the desirable trait could be introduced into a widely used variety or breed of that organism.



Despite some spectacular successes, such as the breeding of high-yielding varieties of wheat, rice, and hybrid corn, such deliberate crossing can be a hit-or-miss affair. Many desirable traits are controlled by multiple genes, and it is hard to predict the results of a cross or to maintain a prized combination as a pure-breeding variety year after year. In sexual reproduction, combinations of desirable genes are quickly separated by meiosis. Furthermore, traditional breeding takes a long time: many plants and animals take years to reach maturity and then can reproduce only once or twice a year—a far cry from the rapid reproduction of bacteria.

Modern recombinant DNA technology has several advantages over traditional methods of breeding (**Figure 18.14**):

- *The ability to identify specific genes.* The development of genetic markers allows breeders to select for specific desirable genes, making the breeding process more precise and rapid.
- *The ability to introduce any gene from any organism into a plant or animal species.* This ability, combined with mutagenesis techniques, vastly expands the range of possible new traits.
- *The ability to generate new organisms quickly.* Manipulating cells in the laboratory and regenerating a whole plant by cloning is much faster than traditional breeding.

Consequently, recombinant DNA technology has found many applications in agriculture (**Table 18.2**). We will describe a few examples to demonstrate the approaches that plant scientists have used to improve crop plants.

PLANTS THAT MAKE THEIR OWN INSECTICIDES Plants are subject to infections by viruses, bacteria, and fungi, but probably the most important crop pests are herbivorous insects. From the locusts of biblical (and modern) times to the cotton boll weevil, insects have continually eaten the crops people grow.

The development of insecticides has improved the situation somewhat, but insecticides have their own problems. Many, including the organophosphates, are relatively nonspecific and kill beneficial insects in the broader ecosystem as well as crop pests. Some even have

18.14 Genetically Modified Plants versus Conventional Plant Breeding

Plant biotechnology offers many potential advantages over conventional breeding.

TABLE 18.2
Agricultural Applications of Biotechnology under Development

PROBLEM	TECHNOLOGY/GENES
Improving the environmental adaptations of plants	Genes for drought tolerance, salt tolerance
Improving nutritional traits	High-lysine seeds; β -carotene in rice
Improving crops after harvest	Delay of fruit ripening; sweeter vegetables
Using plants as bioreactors	Plastics, oils, and drugs produced in plants

toxic effects on other groups of organisms, including people. What's more, many insecticides persist in the environment for a long time.

Some bacteria protect themselves by producing proteins that can kill insects. For example, the bacterium *Bacillus thuringiensis* produces a protein that is toxic to the insect larvae that prey on it. The toxicity of this protein is 80,000 times greater than that of a typical commercial insecticide. When a hapless larva eats the bacteria, the toxin becomes activated and binds specifically to the insect's gut, producing holes and killing the insect. Dried preparations of *B. thuringiensis* have been sold for decades as safe insecticides that break down rapidly in the environment. But the biodegradation of these preparations is their limitation, because it means that the dried bacteria must be applied repeatedly during the growing season.

A more permanent approach is to have the crop plants themselves make the toxin, and this is exactly what plant scientists have done. The toxin gene from *B. thuringiensis* has been isolated, cloned, and extensively modified by the addition of a plant promoter and other regulatory sequences. Transgenic corn, cotton, soybeans, tomatoes, and other crops are now being grown successfully with this added gene. Pesticide usage by farmers growing these transgenic crops is greatly reduced.

CROPS THAT ARE RESISTANT TO HERBICIDES Herbivorous insects are not the only threat to agriculture. Weeds may grow in fields and compete with crop plants for water and soil nutrients. Glyphosate is a widely used and effective herbicide, or weed killer, that works only on plants. It inhibits an enzyme system in the chloroplast that is involved in the synthesis of amino acids. Glyphosate is a broad-spectrum herbicide that kills most weeds, but unfortunately it also kills crop plants. One solution to this problem is to use it to rid a field of weeds before the crop plants start to grow. But, as any gardener knows, when the crop begins to grow, the weeds reappear. If the crop were not affected by the herbicide, the herbicide could be applied to the field at any time.

Scientists have used expression vectors to make plants that synthesize a different form of the target enzyme for glyphosate that is unaffected by the herbicide. The gene for this enzyme has been inserted into corn, cotton, and soybean plants, making them resistant to glyphosate. This technology has expanded rapidly and a large proportion of cotton and soybean plants now carry this gene.

GRAINS WITH IMPROVED NUTRITIONAL CHARACTERISTICS

To remain healthy, humans must consume adequate amounts of β -carotene, which the body converts into vitamin A (see Figure 3.21). About 400 million people worldwide suffer from vitamin A deficiency, which makes them susceptible to infections and blindness. One reason is that rice grains, which do not contain β -carotene, make up a large part of their diets. Other parts of the rice plant, and indeed many plants

and other organisms, contain enzymes for the biochemical pathway that leads to β -carotene production.

Plant biologists Ingo Potrykus and Peter Beyer isolated one of the genes for the β -carotene pathway from the bacterium *Erwinia uredovora* and another from daffodil plants. They added a promoter and other signals for expression in the developing rice grain, and then transformed rice plants with the two genes. The resulting rice plants produce grains that look yellow because of their high β -carotene content. A newer variety with a corn gene replacing the one from daffodils makes even more β -carotene and is golden in color (Figure 18.15). A daily intake of about 150 grams of this cooked rice can supply all the β -carotene a person needs. This new transgenic strain has been crossed with strains adapted for various local environments, in the hope of improving the diets of millions of people.

CROPS THAT ADAPT TO THE ENVIRONMENT Agriculture depends on ecological management—tailoring the environment to the needs of crop plants and animals. A farm field is an unnatural, human-designed system that must be carefully managed to maintain optimal conditions for crop growth. For example, excessive irrigation can cause increases in soil salinity. The Fertile Crescent, the region between the Tigris and Euphrates rivers in the Middle East where agriculture probably originated 10,000 years ago, is no longer fertile. It is now a desert, largely because the soil has a high salt concentration. Few plants can grow on salty soils, partly because of osmotic ef-



18.15 Transgenic Rice Rich in β -Carotene Right and middle: The grains from these transgenic rice strains are colored because they make the pigment β -carotene, which is converted to vitamin A in the human body. Left: Normal rice grains do not contain β -carotene.

fects that result in wilting, and partly because excess salt ions are toxic to plant cells.

Some plants can tolerate salty soils because they have a protein that transports Na^+ ions out of the cytoplasm and into the vacuole, where the ions can accumulate without harming plant growth (see Section 5.3 for a description of the plant vacuole). In many salt-intolerant plants, including *Arabidopsis thaliana*, the gene for this protein exists but is inactive. Recombinant DNA technology has allowed scientists to create active versions of this gene, and to use it to transform crop plants such as rapeseed, wheat, and tomatoes. When this gene was added to tomato plants, they grew in water that was four times as salty as the typical lethal level (**Figure 18.16**). This finding raises the prospect of growing useful crops on what were previously unproductive soils.

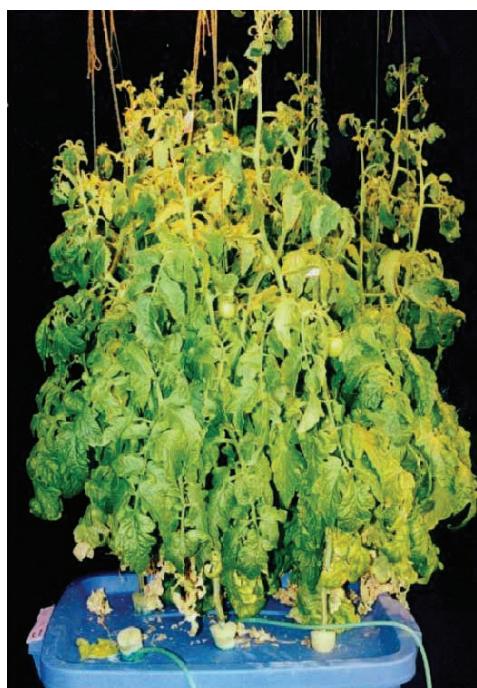
This example illustrates what could become a fundamental shift in the relationship between crop plants and the environment. *Instead of manipulating the environment to suit the plant, biotechnology may allow us to adapt the plant to the environment.* As a result, some of the negative effects of agriculture, such as water pollution, could be lessened.

Biotechnology can be used for environmental cleanup

The thousands of species of bacteria have many unique enzymes and biochemical pathways. Bacteria are nature's recyclers, thriving on many types of nutrients—including what humans refer to as wastes. **Bioremediation** is the use by humans of other organisms to remove contaminants from the environment. Two well-known uses of bacteria for bioremediation are composting and wastewater treatment.

Composting involves the use of bacteria to break down large molecules, including carbon-rich polymers and proteins in waste products such as wood chips, paper, straw, and kitchen

(A)



(B)



scraps. For example, some species of bacteria make cellulase, an enzyme that hydrolyzes cellulose. Bacteria are used in *waste-water treatment* to break down human wastes, paper products, and household chemicals.

Transgenic organisms can also be used to clean up environmental contaminants. As we saw at the opening of this chapter, bacteria are being used to help clean up oil spills. As another example, plants that have been modified to take up heavy metals are being explored as a way to remediate contaminated soils, such as mine tailings (see Section 39.4).

There is public concern about biotechnology

Concerns have been raised about the safety and wisdom of genetically modifying crops and other organisms. These concerns are centered on three claims:

- Genetic manipulation is an unnatural interference with nature.
- Genetically altered foods are unsafe to eat.
- Genetically altered crop plants are dangerous to the environment.

Advocates of biotechnology tend to agree with the first claim. However, they point out that all crops are unnatural in the sense that they come from artificially bred plants growing in a manipulated environment (a farmer's field). Recombinant DNA technology just adds another level of sophistication to these technologies.

To counter the concern about whether genetically engineered crops are safe for human consumption, biotechnology advocates point out that only single genes are added and that these genes are specific for plant function. For example, the *B. thuringiensis* toxin produced by transgenic plants has no effect on people. However, as plant biotechnology moves from adding genes that improve plant growth to adding genes that affect human nutrition, such concerns will become more pressing.

Various negative environmental impacts have been envisaged. There is concern about the possible "escape" of transgenes from crops to other species. If the gene for herbicide resistance, for example, were inadvertently transferred from a crop plant to a closely related weed, that weed could thrive in her-

18.16 Salt-Tolerant Tomato Plants

Transgenic plants containing a gene for salt tolerance thrive in salty water (A), while plants without the transgene die (B). This technology may allow crops to be grown on salty soils.

bicide-treated areas. Another negative impact would be the development of new super-weeds from transgenic crops. For example, a drought tolerant crop plant might spread into, and upset the ecology of, a desert. Or beneficial insects could eat plant materials containing *B. thuringiensis* toxin and die. Transgenic plants undergo extensive field-testing before they are approved for use, but the complexity of the biological world makes it impossible to predict all potential environmental effects of transgenic organisms. In fact, some spreading of transgenes has been detected. Because of the potential benefits of agricultural biotechnology (see Table 18.2), scientists believe that it is wise to proceed with caution.

18.6 RECAP

Biotechnology has been used to produce medicines and to develop transgenic plants with improved agricultural and nutritional characteristics.

- What are the advantages of using biotechnology for plant breeding compared with traditional methods? **See Figure 18.14**
- What are some of the concerns that people might have about agricultural biotechnology? **See pp. 402–403**

CHAPTER SUMMARY

18.1 What Is Recombinant DNA?

- **Recombinant DNA** is formed by the combination of two DNA sequences from different sources. **Review Figure 18.1**
- Many **restriction enzymes** make staggered cuts in the two strands of DNA, creating fragments that have **sticky ends** with unpaired bases.
- DNA fragments with sticky ends can be used to create recombinant DNA. DNA molecules from different sources can be cut with the same restriction enzyme and spliced together using **DNA ligase**. **Review Figure 18.2**

18.2 How Are New Genes Inserted into Cells?

- One goal of recombinant DNA technology is to **clone** a particular gene, either for analysis or to produce its protein product in quantity.
- Bacteria, yeasts, and cultured plant and animal cells are commonly used as hosts for recombinant DNA. The insertion of foreign DNA into host cells is called **transformation** or **transfection** (for animal cells). Transformed or transfected cells are called **transgenic** cells.
- Various methods are used to get recombinant DNA into cells. These include chemical or electrical treatment of the cells, the use of viral vectors, and injection. *Agrobacterium tumefaciens* is often used to insert DNA into plant cells.
- To identify host cells that have taken up a foreign gene, the inserted sequence can be tagged with one or more **reporter genes**, which are genetic markers with easily identifiable phenotypes. **Selectable markers** allow for the selective growth of transgenic cells.
- Replication of the foreign gene in the host cell requires that it become part of a segment of DNA that contains a **replicon** (origin and terminus of replication).
- **Vectors** are DNA sequences that can carry new DNA into host cells. Plasmids and viruses are commonly used as vectors. **Review Figure 18.3**

18.3 What Sources of DNA Are Used in Cloning?

- DNA fragments from a genome can be inserted into host cells to create a **genomic library**. **Review Figure 18.6A**
- The mRNAs produced in a certain tissue at a certain time can be extracted and used to create **complementary DNA (cDNA)** by reverse transcription. **Review Figure 18.6B**

- PCR products can be used for cloning.
- Synthetic DNA containing any desired sequence can be made and mutated in the laboratory.

18.4 What Other Tools Are Used to Study DNA Function?

- Homologous recombination can be used to **knock out** a gene in a living organism. **Review Figure 18.7**
- Gene silencing techniques can be used to inactivate the mRNA transcript of a gene, which may provide clues to the gene's function. Artificially created **antisense RNA** or **siRNA** can be added to a cell to prevent translation of a specific mRNA. **Review Figure 18.8**
- **DNA microarray** technology permits the screening of thousands of cDNA sequences at the same time. **Review Figure 18.9, ANIMATED TUTORIAL 18.1**

18.5 What Is Biotechnology?

- **Biotechnology** is the use of living cells to produce materials useful to people. Recombinant DNA technology has resulted in a boom in biotechnology.
- **Expression vectors** allow a transgene to be expressed in a host cell. **Review Figure 18.11, WEB ACTIVITY 18.1**

18.6 How Is Biotechnology Changing Medicine, Agriculture, and the Environment?

- Recombinant DNA techniques have been used to make medically useful proteins. **Review Figure 18.12**
- **Pharming** is the use of transgenic plants or animals to produce pharmaceuticals. **Review Figure 18.13**
- Because recombinant DNA technology has several advantages over traditional agricultural biotechnology, it is being extensively applied to agriculture. **Review Figure 18.14**
- Transgenic crop plants can be adapted to their environments, rather than vice versa.
- **Bioremediation** is the use of organisms, which are often genetically modified, to improve the environment by breaking down pollutants.
- There is public concern about the application of recombinant DNA technology to food production.

SELF-QUIZ

1. Restriction enzymes
 - a. play no role in bacteria.
 - b. cleave DNA at highly specific recognition sequences.
 - c. are inserted into bacteria by bacteriophage.
 - d. are made only by eukaryotic cells.
 - e. add methyl groups to specific DNA sequences.
2. Which of the following is used as a reporter gene in recombinant DNA work with bacteria as host cells?
 - a. rRNA
 - b. Green fluorescent protein
 - c. Antibiotic sensitivity
 - d. Ability to make ornithine
 - e. Vitamin synthesis
3. From the list below, select the sequence of steps for inserting a piece of foreign DNA into a plasmid vector, introducing the plasmid into bacteria, and verifying that the plasmid and the foreign gene are present:
 - (1) Transform host cells.
 - (2) Select for the lack of plasmid reporter gene 1 function.
 - (3) Select for the plasmid reporter gene 2 function.
 - (4) Digest vector and foreign DNA with a restriction enzyme, which inactivates plasmid reporter gene 1.
 - (5) Ligate the digested plasmid together with the foreign DNA.
 - a. 4, 5, 1, 3, 2
 - b. 4, 5, 1, 2, 3
 - c. 1, 3, 4, 2, 5
 - d. 3, 2, 1, 4, 5
 - e. 1, 3, 2, 5, 4
4. Possession of which feature is *not* desirable in a vector for gene cloning?
 - a. An origin of DNA replication
 - b. Genetic markers for the presence of the vector
 - c. Many recognition sequences for the restriction enzyme to be used
 - d. One recognition sequence each for one or more different restriction enzymes
 - e. Genes other than the target for transfection
5. RNA interference (RNAi) inhibits
 - a. DNA replication.
 - b. neither transcription nor translation of specific genes.
 - c. recognition of the promoter by RNA polymerase.
 - d. transcription of all genes.
 - e. translation of specific mRNAs.
6. Complementary DNA (cDNA)
 - a. is produced from ribonucleoside triphosphates.
 - b. is produced by reverse transcription.
 - c. is the “other strand” of single-stranded DNA in a virus.
 - d. requires no template for its synthesis.
 - e. cannot be placed into a vector because it has the opposite base sequence of the vector DNA.
7. In a genomic library of frog DNA in *E. coli* bacteria,
 - a. all bacterial cells have the same sequences of frog DNA.
 - b. all bacterial cells have different sequences of frog DNA.
 - c. each bacterial cell has a random fragment of frog DNA.
 - d. each bacterial cell has many fragments of frog DNA.
 - e. the frog DNA is transcribed into mRNA in the bacterial cells.
8. An expression vector requires all of the following except
 - a. genes for ribosomal RNA.
 - b. a reporter gene.
 - c. a promoter of transcription.
 - d. an origin of DNA replication.
 - e. restriction enzyme recognition sequences.
9. “Pharming” is a term that describes
 - a. the use of animals in transgenic research.
 - b. plants making genetically altered foods.
 - c. synthesis of recombinant drugs by bacteria.
 - d. large-scale production of cloned animals.
 - e. synthesis of a drug by a transgenic plant or animal.
10. Which of the following could *not* be used to test whether expression of a particular gene is necessary for a particular biological function?
 - a. RNAi
 - b. Knockout technology
 - c. Antisense
 - d. Mutant tRNA
 - e. Transposon mutagenesis

FOR DISCUSSION

1. Compare PCR (see Section 13.5) and cloning as methods to amplify a gene. What are the requirements, benefits, and drawbacks of each method?
2. As specifically as you can, outline the steps you would take to (a) insert and express the gene for a new, nutritious seed protein in wheat, and (b) insert and express a gene for a human enzyme in sheep’s milk.
3. Compare traditional genetic methods with molecular methods for producing genetically altered plants. For each case, describe (a) sources of new genes; (b) numbers of genes transferred; and (c) how long the process takes.

ADDITIONAL INVESTIGATION

Green fluorescent protein (GFP) from a jellyfish can be incorporated into a vector as a reporter gene to signal the presence of the vector in a host cell (see Figure 18.5). How would you alter

the technique in Figure 18.4 to substitute GFP for one (or both) of the antibiotic resistance markers?

WORKING WITH DATA (GO TO yourBioPortal.com)

Recombinant DNA In 1973, Stanley Cohen and Herbert Boyer pioneered the field of recombinant DNA technology when they demonstrated that biologically functional recombinant bacterial plasmids can be constructed in the laboratory

(Figure 18.1). In this exercise, you will examine their original research article and calculations from their data that show that recombinant DNA was made.

19

Differential Gene Expression in Development

On track with stem cells

In horse racing, bettors speak of the “future book” odds on a horse’s chances in an upcoming race. On the morning after winning a race in 2005, the future book odds for Greg’s Gold did not look good—he was limping because of a shredded tendon in his right front leg. A tendon is like a rubber band connecting muscles and bones, and tendons in the legs store energy when an animal runs. Typically, a damaged tendon is allowed to heal naturally, but scar tissue makes it less flexible, and a horse cannot run as fast as it did before injuring a tendon. So it looked as if Greg’s Gold might have to retire from racing.

Greg’s Gold’s trainer, David Hofmans, decided to try a new therapy. A veterinarian removed a small amount of adipose (fatty) tissue from the horse’s hindquarters and sent it to a cell biology laboratory. There, the tissue was treated with enzymes to digest the extracellular molecules that held the cells together. Several cell populations were obtained, among them mesenchymal stem cells.

Stem cells are actively dividing, unspecialized cells that have the potential to produce different cell types depending on the signals they receive from the body. Mesenchymal stem cells are able to differentiate into various kinds of connective tissue, including bone, cartilage, blood vessels, tendons, and muscle.

Two days after the tissue was taken, Greg’s Gold’s veterinarian received the stem cells back from the lab and injected them into the site of the damaged tendon. After several months, the tendon healed with little scar tissue, and Greg’s Gold’s trainer returned him to the racetrack. Greg’s Gold raced for almost two more years, winning over \$1 million in purse money before being retired.

The mesenchymal stem cell treatment has been used successfully on several thousand horses, and on dogs with arthritis. Most stem cell therapies for humans are still at the experimental stage, particularly in the United States,

where controversy over the use of embryonic stem cells has slowed the progress of research and the adoption of therapeutic techniques. But in Japan, women undergoing reconstructive surgery after the removal of breast cancer have had more favorable outcomes when treated with their own mesenchymal stem cells. Bone marrow transplantation is one form of stem cell therapy that has been used successfully for more than thirty years in the United States, to treat patients with cancers such as leukemia and lymphoma.



Greg's Gold Fat stem cells helped repair damage to his tendons and he was able to race—and win—again.



Fat as a Source of Stem Cells This centrifuge separates dense fatty tissues from the lighter stem cells. Stem cells from fat have been found to be capable of differentiating into several specialized cell types.

The processes by which an unspecialized stem cell proliferates and forms specialized cells and tissues with distinctive appearances and functions are similar to the developmental processes that occur in the embryo. Much of our knowledge of developmental biology has come from studies on model organisms such as the fruit fly *Drosophila melanogaster*, the nematode worm *Caenorhabditis elegans*, zebrafish, the mouse, and the small flowering plant *Arabidopsis thaliana*. Eukaryotes share many similar genes, and the cellular and molecular principles underlying their development also turn out to be similar. Thus discoveries from one organism can aid us in understanding other organisms, including ourselves.

IN THIS CHAPTER we begin by describing how almost every cell in a multicellular organism contains all of the genes present in the zygote that gave rise to that organism. Then we explain how cellular changes during development result from the differential expression of those genes. Finally, we show how the various mechanisms of transcriptional control and chemical signaling that are discussed in previous chapters work together to produce a complex organism.

CHAPTER OUTLINE

- 19.1 What Are the Processes of Development?
- 19.2 Is Cell Differentiation Irreversible?
- 19.3 What Is the Role of Gene Expression in Cell Differentiation?
- 19.4 How Is Cell Fate Determined?
- 19.5 How Does Gene Expression Determine Pattern Formation?

19.1 What Are the Processes of Development?

Development is the process by which a multicellular organism, beginning with a single cell, goes through a series of changes, taking on the successive forms that characterize its life cycle (**Figure 19.1**). After the egg is fertilized, it is called a zygote, and in the earliest stages of development a plant or animal is called an **embryo**. Sometimes the embryo is contained within a protective structure such as a seed coat, an eggshell, or a uterus. An embryo does not photosynthesize or feed itself. Instead, it obtains its food from its mother either directly (via the placenta) or indirectly (by way of nutrients stored in a seed or egg). A series of embryonic stages precedes the birth of the new, independent organism. Many organisms continue to develop throughout their life cycle, with development ceasing only with death.

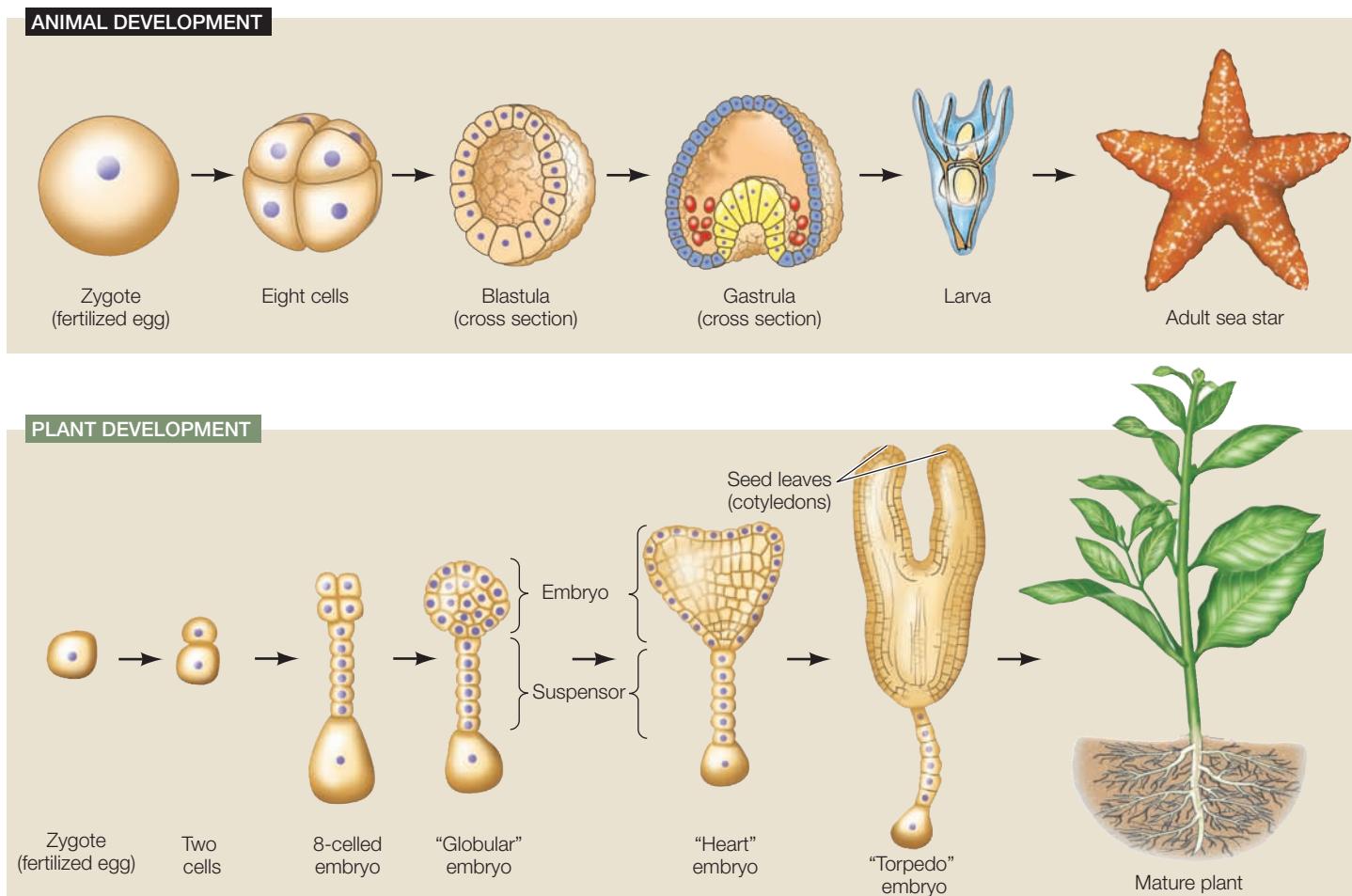
Development involves distinct but overlapping processes

The developmental changes an organism undergoes as it progresses from an embryo to mature adulthood involve four processes:

- **Determination** sets the developmental *fate* of a cell—what type of cell it will become—even before any characteristics of that cell type are observable. For example, the mesenchymal stem cells described in the opening story look unspecialized, but their fate to become connective tissue cells has already been determined.
- **Differentiation** is the process by which different types of cells arise, leading to cells with specific structures and functions. For example, mesenchymal stem cells differentiate to become muscle, fat, tendon, or other connective tissue cells.
- **Morphogenesis** (Greek for “origin of form”) is the organization and spatial distribution of differentiated cells into the multicellular body and its organs.
- **Growth** is the increase in size of the body and its organs by cell division and cell expansion.

Determination and differentiation occur largely because of differential gene expression. The cells that arise from repeated mitoses in the early embryo may look the same superficially, but they soon begin to differ in terms of which of the thousands of genes in the genome are expressed.

Morphogenesis involves differential gene expression and the interplay of signals between cells. Morphogenesis can occur in several ways:



19.1 From Fertilized Egg to Adult The stages of development from zygote to maturity are shown for an animal and for a plant. The blastula is a hollow sphere of cells; the gastrula has three cell layers.

yourBioPortal.com
GO TO Web Activity 19.1 • Stages of Development

- Cell division is important in both plants and animals.
- Cell expansion is especially important in plant development, where a cell's position and shape are constrained by the cell wall.
- Cell movements are very important in animal morphogenesis (see Section 44.2)
- Apoptosis (programmed cell death) is essential in organ development.

Growth can occur by an increase in the number of cells or by the enlargement of existing cells. Growth continues throughout the individual's life in some organisms, but reaches a more or less stable end point in others.

Cell fates become progressively more restricted during development

During development, each undifferentiated cell will become part of a particular type of tissue—this is referred to as the **cell fate** of that undifferentiated cell. A cell's fate is a function of both differential gene expression and morphogenesis. The role of

morphogenesis in determining cell fate was revealed in experiments in which undifferentiated cells were removed from specific locations in early embryos and grafted into new positions on other embryos. The cells were marked with stains so that their development into adult structures could be traced. Such experiments on amphibian embryos indicated that the fates of early embryonic cells are not irrevocably determined, but depend on the cells' environment and stage of development (Figure 19.2). In this example, the cells that would have become skin tissue if left in place became brain or notochord tissues, depending on the locations of the grafts.

But as development proceeds from zygote to mature organism, the developmental potential of cells becomes more restricted. For example, if tissue is removed from the brain area of a later-stage frog embryo, it will become brain tissue, even if transplanted to a part of an early-stage embryo that is destined to become another structure.

As we will discuss in this chapter, cell fate determination is influenced by changes in gene expression as well as the extracellular environment. Determination is not something that is visible under the microscope—cells do not change their appearance when they become determined. Determination is followed by differentiation—the actual changes in biochemistry, structure, and function that result in cells of different types. *Determination is a commitment; the final realization of that commitment is differentiation.*

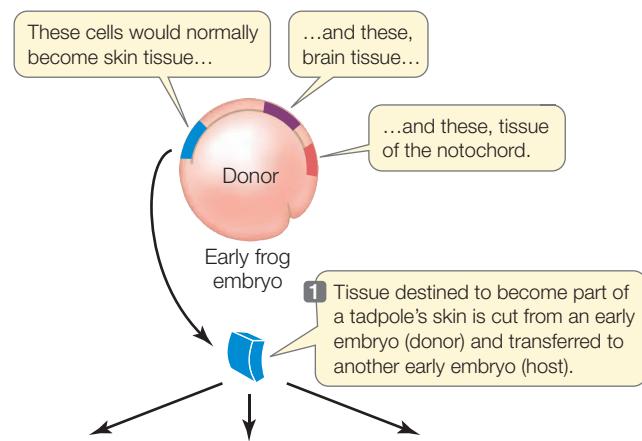
INVESTIGATING LIFE

19.2 Developmental Potential in Early Frog Embryos

In an early embryo, the cells look alike. But marking experiments suggested that the fates of these cells were determined early in development. Was the fate of a cell irrevocable or did it still retain the ability to become a different cell type? To answer this question, biologists transplanted cells from one location in one embryo to a different location in a second embryo. The cells took on the fate of cells at the new location. Therefore, cells in the early embryo retained the ability to form other cell types if placed in the right environment.

HYPOTHESIS The fate of the cells in an early amphibian embryo is irrevocably determined.

METHOD



Experiment 1

2a The donor tissue is transplanted to the "brain" region.



Experiment 2

2b The donor tissue is transplanted to the "notochord" region.

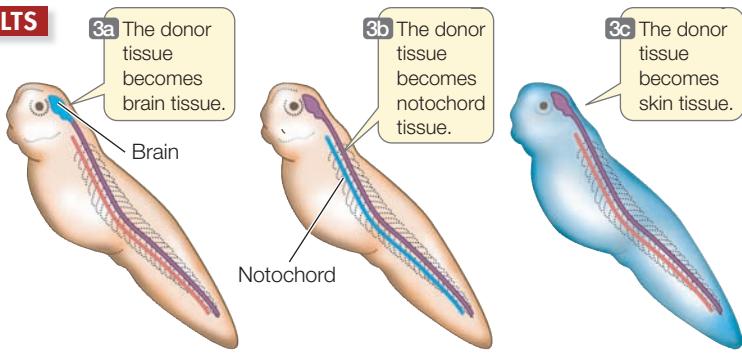


Experiment 3

2c The donor tissue is transplanted to the "skin" region (control).



RESULTS



CONCLUSION The hypothesis is rejected. Cell fates in the early embryo are not determined, but can change depending on the environment.

FURTHER INVESTIGATION: What would happen if tissue from an adult were transplanted into an early embryo?

Go to yourBioPortal.com for original citations, discussions, and relevant links for all INVESTIGATING LIFE figures.

19.1 RECAP

Development takes place via the processes of determination, differentiation, morphogenesis, and growth. Cells in the very early embryo have not yet had their fates determined; as development proceeds, their potential fates become more and more restricted.

- What are the four processes of development? See p. 405
- Explain what the experiment in Figure 19.2 told us about how cell fates become determined. See p. 407

Is a mesophyll cell in a plant leaf or a liver cell in a human being irrevocably committed to that specialization? Under the right experimental circumstances, differentiation is reversible in some cells. The next section describes how the genomes of some cells can be induced to express different sets of genes used in differentiation.

19.2 Is Cell Differentiation Irreversible?

A zygote has the ability to give rise to every type of cell in the adult body; in other words, it is **totipotent** (*toti*, "all"; *potent*, "capable"). Its genome contains instructions for all of the structures and functions that will arise throughout the life cycle of the organism. Later in development, the cellular descendants of the zygote lose their totipotency and become determined. These determined cells then differentiate into specialized cells. The human liver cell and the leaf mesophyll cell generally retain their differentiated forms and functions throughout their lives. But this does not necessarily mean that they have irrevocably lost their totipotency. Most of the differentiated cells of an animal or plant have nuclei containing the entire genome of the organism and therefore have the genetic capacity for totipotency. We explore here several examples of how this capacity has been demonstrated experimentally.

Plant cells can be totipotent

A carrot root cell normally faces a dark future. It cannot photosynthesize and generally does not give rise to new carrot plants. However, in 1958 Frederick Steward at Cornell University showed that if he isolated cells from a carrot

root and maintained them in a suitable nutrient medium, he could induce them to dedifferentiate—to lose their differentiated characteristics. The cells could divide and give rise to masses of undifferentiated cells called *calli* (singular *callus*), which could be maintained in culture indefinitely. But, if they were provided with the right chemical cues, the cells could develop into embryos and eventually into complete new plants (Figure 19.3). Since the new plants were genetically identical to the cells from which they came, they were clones of the original carrot plant.

The ability to clone an entire carrot plant from a differentiated root cell indicated that the cell contained the entire carrot genome, and that under the right conditions, the cell and its descendants could express the appropriate genes in the right sequence to form a new plant. Many types of cells from other plant species show similar behavior in the laboratory. This ability to generate a whole plant from a single cell has been invaluable in agriculture and forestry. For example, trees from planted forests are used in making paper, lumber, and other products. To replace the trees reliably, forestry companies regenerate new trees from the leaves of selected trees with desirable traits. The characteristics of these clones are more uniform and predictable than those of trees grown from seeds.

Nuclear transfer allows the cloning of animals

Animal somatic cells cannot be manipulated as easily as plant cells can. However, experiments such as the one shown in Figure 19.2 have demonstrated the totipotency of early embryonic cells from animals. In humans, this totipotency permits both genetic screening (see Section 15.5) and certain assisted reproductive technologies (see Section 43.4). A human embryo can be isolated in the laboratory and one or a few cells removed and examined to determine whether a certain genetic condition is present. Due to their totipotency, the remaining cells can develop into a complete embryo, which can be implanted into the mother's uterus, where it develops into a normal fetus and infant.

Until recently, it was not possible to induce a cell from a fully developed animal to dedifferentiate and then redifferentiate into another cell type. However, nuclear transfer experiments have shown that the genetic information from an animal cell can be used to create cloned animals. Robert Briggs and Thomas King performed the first such experiments in the 1950s using frog embryos. First they removed the nucleus from an unfertilized egg, forming an *enucleated* egg. Then, with a very fine glass needle, they punctured a cell from an early embryo and drew up part of its contents, including the nucleus, which they injected into the enucleated egg. They stimulated the eggs to divide, and many went on to form embryos, and eventually frogs, that were clones from the original implanted nucleus. These experiments led to two important conclusions:

- No information is lost from the nuclei of cells as they pass through the early stages of embryonic development. This fundamental principle of developmental biology is known as **genomic equivalence**.
- The cytoplasmic environment around a cell nucleus can modify its fate.

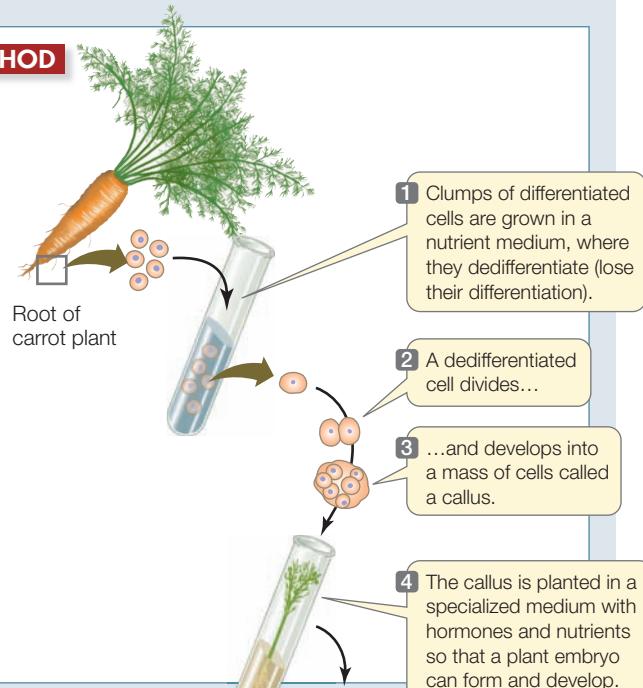
INVESTIGATING LIFE

19.3 Cloning a Plant

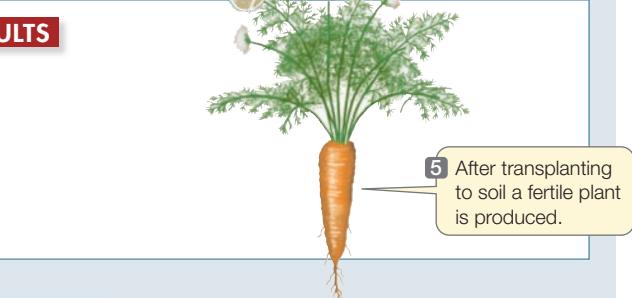
When cells were removed from a plant and put into a medium with nutrients and hormones, they lost many of their specialized features—in other words, they dedifferentiated. Did these cells retain the ability to differentiate again? Frederick Steward found that a cultured carrot cell did indeed retain the ability to develop into an embryo and a new plant.

HYPOTHESIS Differentiated plant cells can be totipotent and can be induced to generate all types of the plant's cells.

METHOD



RESULTS



CONCLUSION Differentiated plant cells can be totipotent.

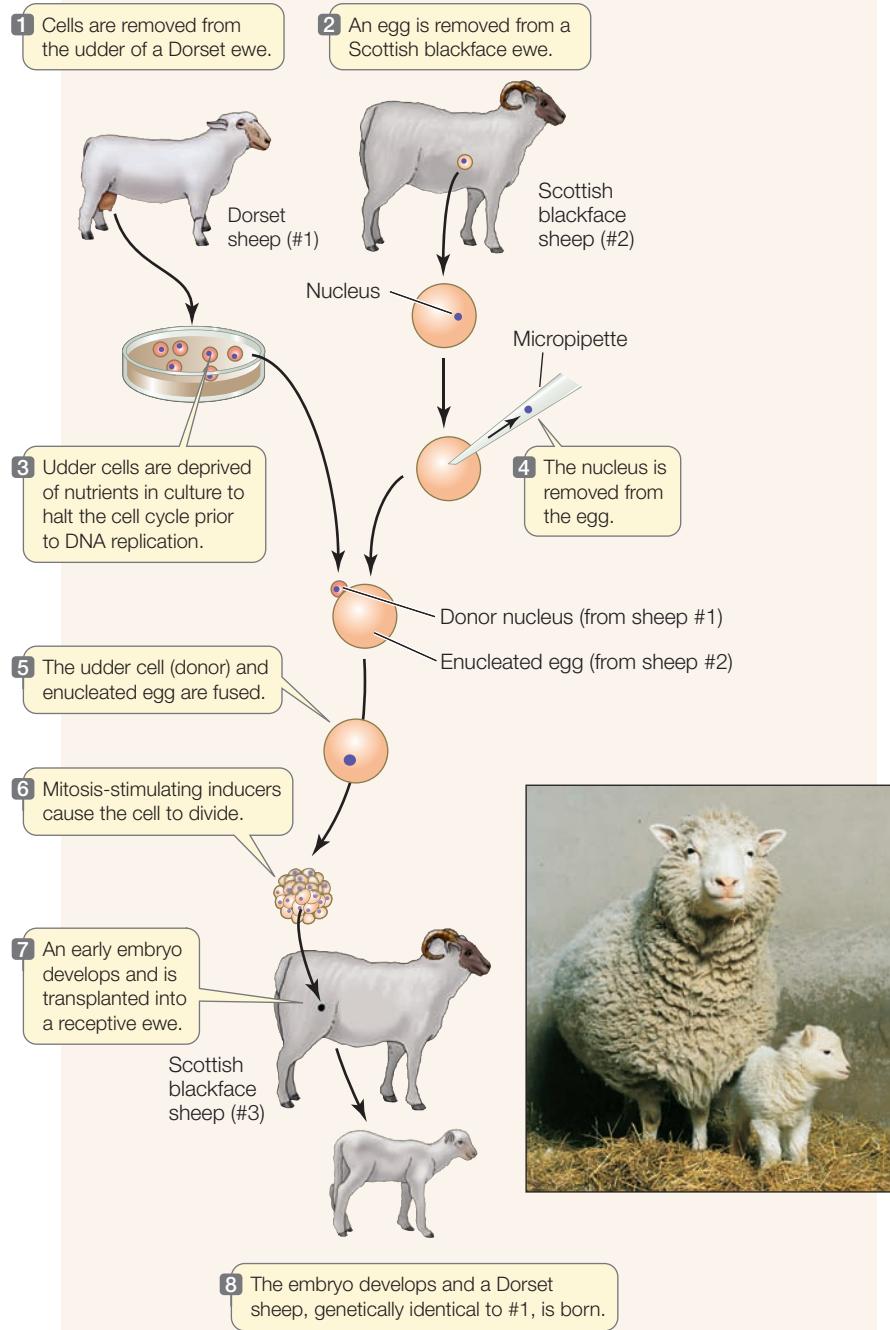
Go to yourBioPortal.com for original citations, discussions, and relevant links for all INVESTIGATING LIFE figures.

In 1996, Ian Wilmut and his colleagues in Scotland cloned the first mammal by the cell fusion method. To produce donor cells suitable for nuclear transfer, they took differentiated cells from a ewe's udder and starved them of nutrients for a week, halting the cells in the G1 phase of the cell cycle. One of these cells was fused with an enucleated egg from a different breed

TOOLS FOR INVESTIGATING LIFE

19.4 Cloning a Mammal

The experimental procedure described here produced the first cloned mammal, a Dorset sheep named Dolly (shown on the left in the photo). As an adult, Dolly mated and subsequently gave birth to a normal offspring (the lamb on the right), thus proving the genetic viability of cloned mammals.



of ewe. Signals from the egg's cytoplasm stimulated the donor nucleus to enter S phase, and the rest of the cell cycle proceeded normally. After several cell divisions, the resulting early embryo was transplanted into the womb of a surrogate mother (**Figure 19.4**).

Out of 277 successful attempts to fuse adult cells with enucleated eggs, one lamb survived to be born; she was named Dolly, and she became world-famous overnight. DNA analyses confirmed that Dolly's nuclear genes were identical to those of the ewe from whose udder the donor nucleus had been obtained. Dolly grew to adulthood, mated, and produced offspring in the normal manner, thus proving her status as a fully functioning adult animal.

Many other animal species, including cats, dogs, horses, pigs, rabbits, and mice have since been cloned by nuclear transfer. The cloning of animals has practical uses and has given us important information about developmental biology. There are several reasons to clone animals:

- *Expansion of the numbers of valuable animals:* One goal of Wilmut's experiments was to develop a method of cloning transgenic animals carrying genes with therapeutic properties. For example, a cow that was genetically engineered to make human growth hormone in milk has been cloned to produce two more cows that do the same thing. Only 15 such cows would supply the world's need for this medication, which is used to treat short stature due to growth hormone deficiency.
- *Preservation of endangered species:* The banteng, a relative of the cow, was the first endangered animal to be cloned, using a cow enucleated egg and a cow surrogate mother. Cloning may be the only way to save endangered species with low rates of natural reproduction, such as the giant panda.
- *Preservation of pets:* Many people get great personal benefit from pets, and the death of a pet can be devastating. Companies have been set up to clone cats and dogs from cells provided by their owners. Of course, the behavioral characteristics of the beloved pet, which are certainly derived in part from the environment, may not be the same in the cloned pet as in its genetic parent.

Multipotent stem cells differentiate in response to environmental signals

In plants, the growing regions at the tips of the roots and stems contain **meristems**, which are clusters of undifferentiated, rapidly dividing **stem cells**. These cells can give rise to the specialized cell types that make up the various parts of roots and stems. In general, plants have far fewer (15–20) broad cell types than animals (as many as 200).

In mammals, stem cells are found in adult tissues that need frequent cell replacement, such as the skin, the inner lining of the intestine, and the bone marrow, where blood and other types

of cells are formed. Canadian cell biologists Ernest McCulloch and James Till discovered mammalian stem cells in the early 1960s when they injected bone marrow cells into adult mice. They noticed that the recipient mice developed small clumps of tissue in the spleen. When they looked more carefully at the clumps, they found that each was composed of undifferentiated stem cells. Before this, stem cells were believed to be present only in animal embryos.

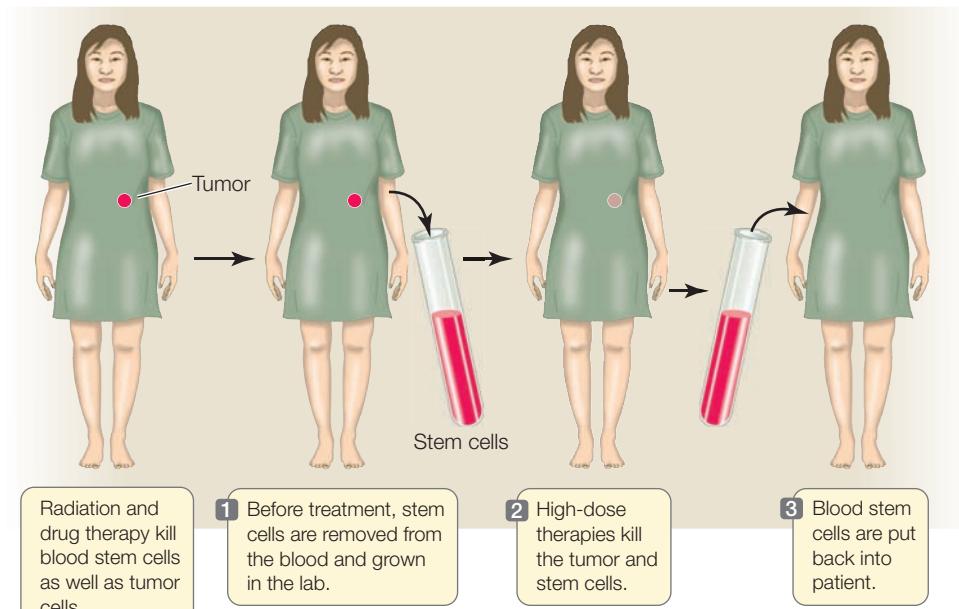
As they divide, stem cells produce daughter cells that differentiate to replace dead cells and maintain the tissues. These adult stem cells in animals are not totipotent, because their ability to differentiate is limited to a relatively few cell types. In other words, they are **multipotent**. For example, there are two types of multipotent stem cells in bone marrow. One type (called hematopoietic stem cells) produces the various kinds of red and white blood cells, while the other type (mesenchymal stem cells) produces the cells that make bone and surrounding *connective* tissues, such as muscle.

The differentiation of multipotent stem cells is “on demand.” The blood cells that differentiate in the bone marrow do so in response to specific signals known as growth factors. This is the basis of an important cancer therapy called *hematopoietic stem cell transplantation* (HSCT) (Figure 19.5). Because some treatments that kill cancer cells also kill other dividing cells, bone marrow stem cells in patients will die if exposed to these treatments. To circumvent this problem, stem cells are removed from the patient’s blood and given growth factors to increase their numbers in the laboratory. The cells are stored during treatment, and then added back to populate the depleted bone marrow when treatment is over. The stored stem cells retain their ability to differentiate in the bone marrow environment. By allowing the use of high doses of treatment to kill tumors, bone marrow transplantation saves thousands of lives each year.

Adjacent cells can also influence stem cell differentiation. We saw this in the opening story of this chapter, in which stem cells from fat differentiated to form cells of the tendon. Bone marrow stem cells that can form muscle will do so if implanted into the heart. Such stem cell transplantation for heart repair has been demonstrated in animals and even in people who had heart attacks, in experiments that used the stem cells to repair a damaged heart. Multipotent stem cells have been found in many organs and tissues, and their use in treating diseases is under intensive investigation.

Pluripotent stem cells can be obtained in two ways

As stated earlier, totipotent stem cells that can form an entire new animal are found only in very early embryos. In both mice and humans, the earliest embryonic stage before differentiation occurs is called a *blastocyst* (see Figure 44.4). Although they



19.5 Stem Cell Transplantation Multipotent blood stem cells can be used in hematopoietic stem cell transplantation, to replace stem cells destroyed by cancer therapy.

cannot form an entire embryo, a group of cells in the blastocyst still retains the ability to form all of the cells in the body: these cells are **pluripotent** (“pluri,” many; “potent,” capable). In mice, these **embryonic stem cells (ESCs)** can be removed from the blastocyst and grown in laboratory culture almost indefinitely if provided with the right conditions. When cultured mouse ESCs are injected back into a mouse blastocyst, the stem cells mix with the resident cells and differentiate to form all the cell types in the mouse. This indicates that the ESCs do not lose any of their developmental potential while growing in the laboratory.

ESCs growing in the laboratory can also be induced to differentiate in a particular way if the right signal is provided (Figure 19.6A). For example, treatment of mouse ESCs with a derivative of vitamin A causes them to form neurons, while other growth factors induce them to form blood cells. Such experiments demonstrate both the cells’ developmental potential and the roles of environmental signals. This finding raises the possibility of using ESC cultures as sources of differentiated cells to repair specific tissues, such as a damaged pancreas in diabetes, or a brain that malfunctions in Parkinson’s disease.

ESCs can be harvested from human embryos conceived by *in vitro* (“under glass”—in the laboratory) fertilization, with the consent of the donors. Since more than one embryo is usually conceived in this procedure, embryos not used for reproduction might be available for embryonic stem cell isolation. These cells could then be grown in the laboratory and used as sources of tissues for transplantation into patients with tissue damage. There are two problems with this approach:

- Some people object to the destruction of human embryos for this purpose.
- The stem cells, and tissues derived from them, would provoke an immune response in a recipient (see Chapter 42).

19.6 Two Ways to Obtain Pluripotent Stem Cells

Cells Pluripotent stem cells can be obtained either from human embryos (A) or by adding highly expressed genes to skin cells to transform them into stem cells (B).

yourBioPortal.com

GO TO Animated Tutorial 19.1 • Embryonic Stem Cells

Shinya Yamanaka and coworkers at Kyoto University in Japan have developed another way to produce pluripotent stem cells that gets around these two problems (Figure 19.6B). Instead of extracting ESCs from blastocysts, they make **induced pluripotent stem cells (iPS cells)** from skin cells. They developed this method systematically:

1. First, they used gene chips to compare the genes expressed in ESCs with nonstem cells (see Figure 18.9). They found several genes that were uniquely expressed at high levels in ESCs. These genes were believed to be essential to the undifferentiated state and function of stem cells.
2. Next, they isolated the genes and inserted them into a vector for genetic transformation of skin cells (see Section 18.5). They found that the skin cells now expressed the newly added genes at high levels.
3. Finally, they showed that the iPS cells were pluripotent and could be induced to differentiate into many tissues.

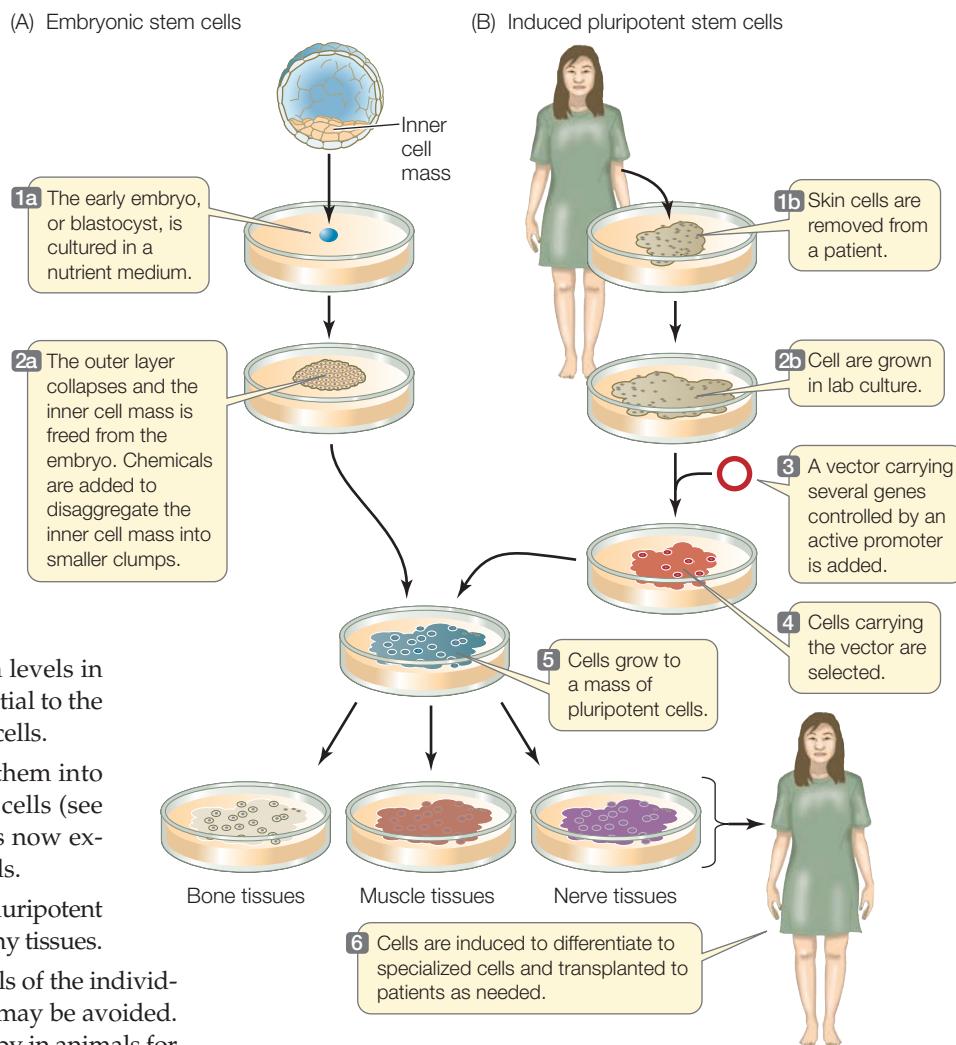
Because the iPS cells can be made from skin cells of the individual who is to be treated, an immune response may be avoided. Such cells have already been used for cell therapy in animals for diseases similar to human Parkinson's disease (a brain disorder), diabetes, and sickle cell anemia. Human uses are sure to follow.

19.2 RECAP

Even differentiated cells retain their ability to differentiate into other cell types, given appropriate chemical signals. This has made cloning and stem cell technologies possible.

- Describe the differences between totipotent, pluripotent, and multipotent cells. **See pp. 408–411**
- How are stem cells found in adult body tissues different from embryonic stem cells? **See pp. 410–411**
- What are the two ways to produce pluripotent stem cells? **See pp. 411–412 and Figure 19.6**

Cloning experiments and observations of stem cells have shown that most differentiated cells in an organism share the same genes. But not all genes are expressed in every cell. What turns gene expression on and off as cells differentiate? In the next section we explore several of the mechanisms controlling the changes in gene expression that lead to cell differentiation.



19.3 What Is the Role of Gene Expression in Cell Differentiation?

Although every cell contains all the genes needed to produce every protein encoded by its genome, each cell expresses only selected genes. For example, certain cells in hair follicles produce keratin, the protein that makes up hair, while other cell types in the body do not. What determines whether a cell will produce keratin? Chapter 16 describes a number of ways in which cells regulate gene expression and protein production—by controlling transcription, translation, and posttranslational protein modifications. But the mechanisms that control gene expression resulting in cell differentiation generally work at the level of transcription.

Differential gene transcription is a hallmark of cell differentiation

The gene for β -globin, one of the protein components of hemoglobin, is expressed in red blood cells as they form in the bone marrow of mammals. That this same gene is also present—but unexpressed—in neurons in the brain (which do not make hemoglobin) can be demonstrated by nucleic acid hybridiza-

tion. Recall that in nucleic acid hybridization, a probe made of single-stranded DNA or RNA of known sequence is added to denatured DNA to reveal complementary coding regions on the DNA template strand (see Figure 15.16). A probe for the β -globin gene can be applied to DNA from brain cells and immature red blood cells (recall that mature mammalian red blood cells lose their nuclei during development). In both cases, the probe finds its complement, showing that the β -globin gene is present in both types of cells. On the other hand, if the β -globin probe is applied to mRNA, rather than DNA, from the two cell types, it finds β -globin mRNA only in the red blood cells, not in the brain cells. This result shows that the gene is expressed in only one of the two cell types.

What leads to this differential gene expression? One well-studied example of cell differentiation is the conversion of undifferentiated muscle precursor cells into cells that are destined to form muscle (Figure 19.7). In the vertebrate embryo these cells come from a layer called the *mesoderm* (see Section 44.2). A key event in the commitment of these cells to become muscle is that they stop dividing. Indeed, in many parts of the embryo, *cell division and cell differentiation are mutually exclusive*. Cell signaling activates the gene for a transcription factor called **MyoD**

(*myoblast-determining gene*). Recall that transcription factors are DNA binding proteins that regulate the expression of specific genes. In this case, MyoD activates the gene for p21, an inhibitor of cyclin-dependent kinases that normally stimulate the cell cycle at G1 (see Figure 11.6). Expression of the *p21* gene causes the cell cycle to stop, and other transcription factors then enter the picture so that differentiation can begin. Interestingly, *myoD* is also activated in stem cells that are present in adult muscle, indicating a role of this transcription factor in repair of muscle as it gets damaged and worn out.

Genes such as *myoD* that direct the most fundamental decisions in development (often by regulating other genes on other chromosomes) usually encode transcription factors. In some cases, a single transcription factor can cause a cell to differentiate in a certain way. In others, complex interactions between genes and proteins determine a sequence of transcriptional events that leads to differential gene expression.

19.3 RECAP

Differentiation involves selective gene expression, controlled at the level of transcription by transcription factors.

- What techniques could you use to identify genes expressed during cell differentiation? **See pp. 412–413**
- What is the role of transcription factors in controlling differentiation? **See p. 413 and Figure 19.7**

Cell differentiation involves extensive transcriptional regulation of genes. But what causes a cell to express one set of genes, and not some other set? In other words, how is a cell's fate determined?

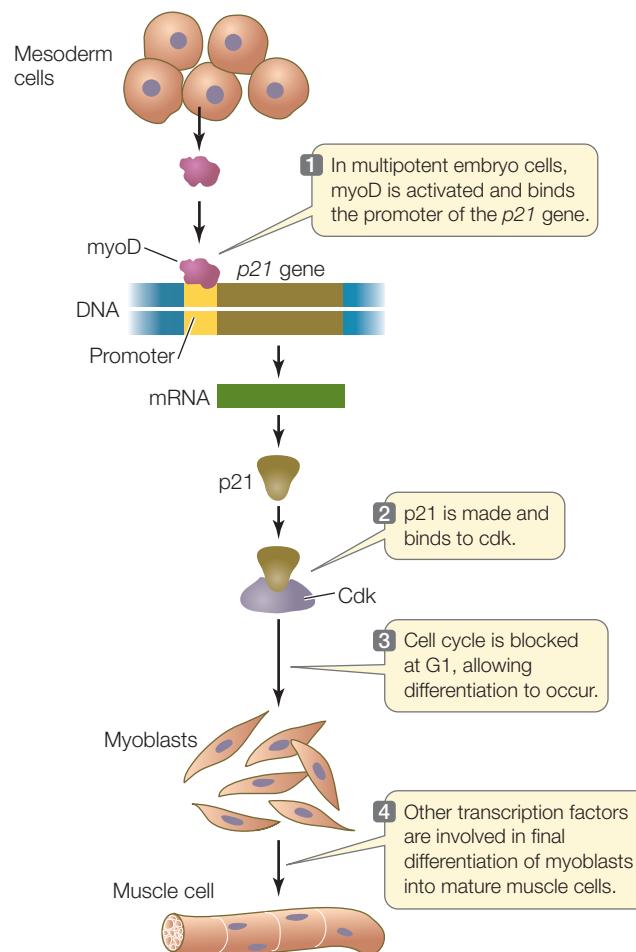
19.4 How Is Cell Fate Determined?

The fertilized egg undergoes many cell divisions to produce the many differentiated cells in the body (such as liver, muscle, and nerve cells). How can one cell produce so many different cell types? There are two ways that this occurs:

- **Cytoplasmic segregation** (unequal cytokinesis). A factor within an egg, zygote, or precursor cell may be unequally distributed in the cytoplasm. After cell division, the factor ends up in some daughter cells or regions of cells, but not others.
- **Induction** (cell-to-cell communication). A factor is actively produced and secreted by certain cells to induce other cells to become determined.

Cytoplasmic segregation can determine polarity and cell fate

Some differences in gene expression patterns are the result of *cytoplasmic* differences between cells. One such cytoplasmic difference is the emergence of distinct “top” and “bottom” ends of an organism or structure; such a difference is called **polarity**.



19.7 Transcription and Differentiation in the Formation of Muscle Cells Activation of a transcription factor, MyoD, is important in muscle cell differentiation.

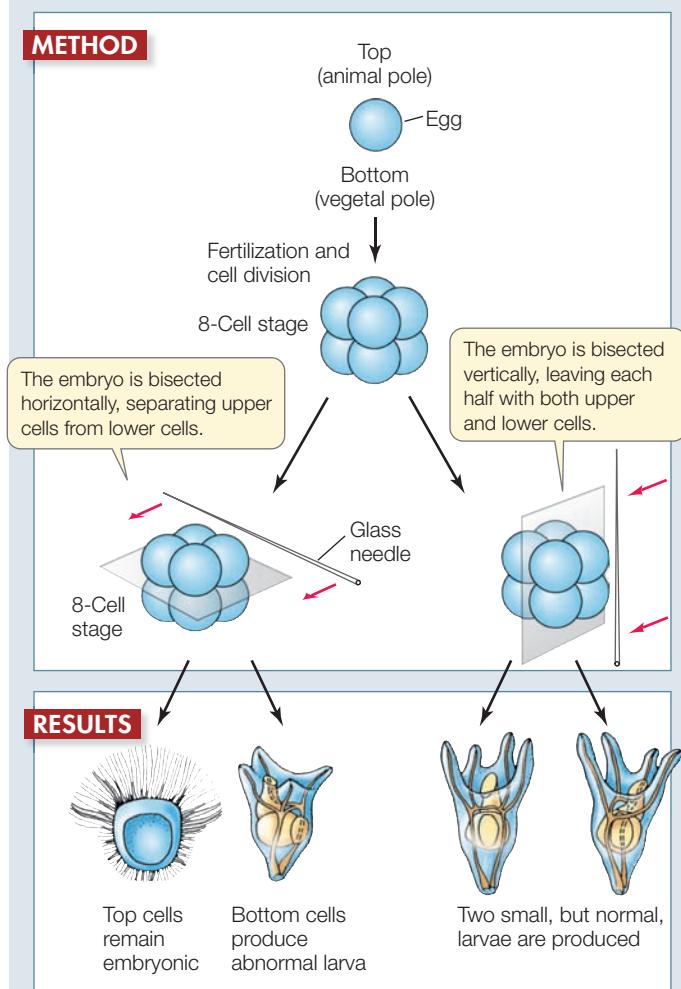
INVESTIGATING LIFE

19.8 Asymmetry in the Early Sea Urchin Embryo

As an embryo develops, cells become determined and their ultimate fate gets more and more narrowly defined. The cells of an eight-celled sea urchin embryo look identical and so might be expected to have the same developmental potential. But do they? Hans Driesch separated different parts of this tiny embryo from one another, to examine their developmental potentials. His experiments showed that even at the eight-cell stage, cell fate determination is underway.

HYPOTHESIS Different regions in the fertilized egg and the embryo have different developmental fates.

METHOD



CONCLUSION The upper and lower halves of a sea urchin embryo differ in their developmental potential.

Go to yourBioPortal.com for original citations, discussions, and relevant links for all INVESTIGATING LIFE figures.

yourBioPortal.com

GO TO Animated Tutorial 19.2 • Early Asymmetry in the Embryo

Many examples of polarity are observed as development proceeds. Our heads are distinct from our rear ends, and the distal (far) ends of our arms and legs (wrists, ankles, fingers, toes)

differ from the proximal (near) ends (shoulders and hips). Polarity may develop early; even within the fertilized egg, the yolk and other factors are often distributed asymmetrically. During early development in animals, polarity is specified by an *animal pole* at the top of the zygote and a *vegetal pole* at the bottom.

In the early twentieth century Hans Driesch at the Marine Biological Station in Naples, Italy, demonstrated the effects of cytoplasmic segregation on development (Figure 19.8). Very early development in sea urchins occurs by rapid, equal mitotic divisions of the fertilized egg; there is no increase in size at this stage, just a partitioning of the cells. If an eight-cell embryo is carefully separated vertically into two four-celled halves, both halves develop into normal (albeit small) larvae. But if an eight-cell embryo is cut horizontally, the top half does not develop at all, while the bottom half develops into a small, abnormal larva.

Clearly, then, there must be at least one factor essential for development that is segregated in the vegetal half of the sea urchin egg, such that the bottom cells of the 8-cell embryo get this essential factor and the top cells do not. Experiments have established that certain materials, called **cytoplasmic determinants**, are distributed unequally in the egg cytoplasm. Cytoplasmic determinants play roles in directing the embryonic development of many organisms (Figure 19.9). What are these determinants and what accounts for their unequal distribution?

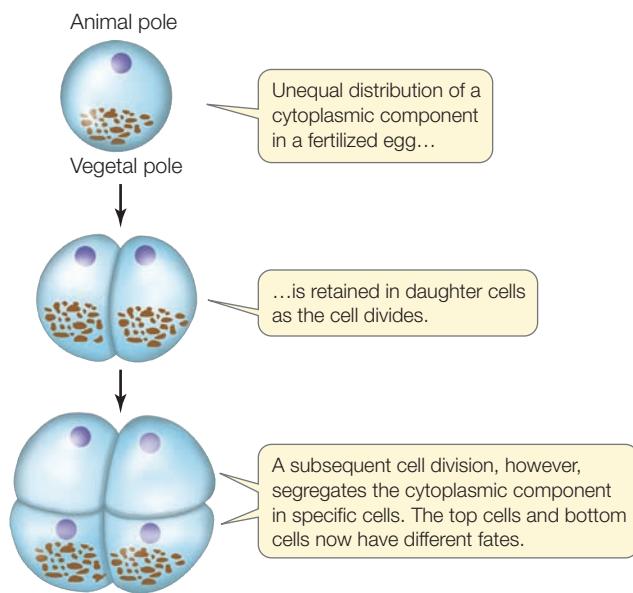
The cytoskeleton contributes to the asymmetric distribution of cytoplasmic determinants in the egg. Recall from Section 5.3 that an important function of the microtubules and microfilaments in the cytoskeleton is to help move materials in the cell. Two properties allow these structures to accomplish this:

- Microtubules and microfilaments have polarity—they grow by adding subunits to the plus end.
- Cytoskeletal elements can bind specific proteins, which can be used in the transport of mRNA.

For example, in the sea urchin egg, a protein binds to both the growing (+) end of a microfilament and to an mRNA encoding a cytoplasmic determinant. As the microfilament grows toward one end of the cell, it carries the mRNA along with it. The asymmetrical distribution of the mRNA leads to a similar distribution of the protein it encodes. So what were once unspecified cytoplasmic determinants can now be defined in terms of cellular structures, mRNAs, and proteins.

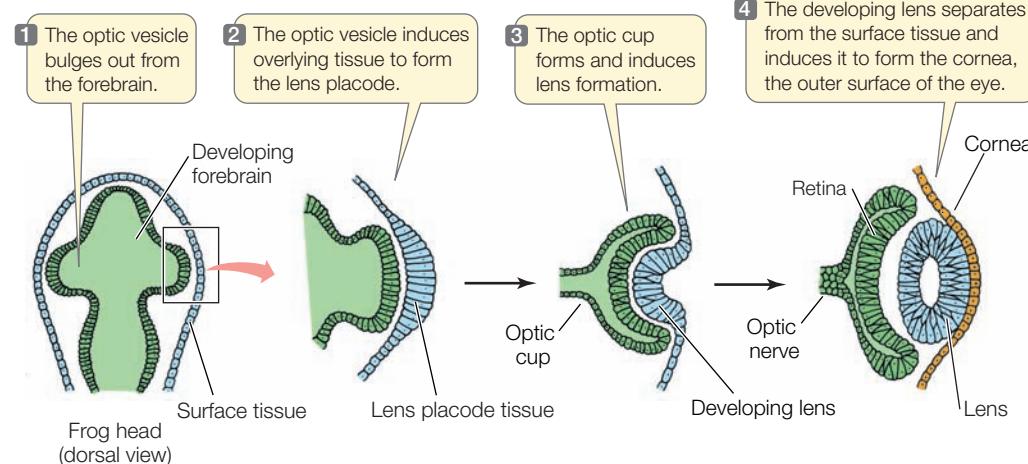
Inducers passing from one cell to another can determine cell fates

The term “induction” has different meanings in different contexts. In biology it can be used broadly to refer to the initiation of, or cause of, a change or process. But in the context of cellular differentiation, it refers to the signaling events by which cells in a developing embryo communicate and influence one another’s developmental fate. Induction involves chemical signals and signal transduction mechanisms. We will describe two examples of this form of induction: one in the developing vertebrate eye, and the other in a developing reproductive structure of the nematode *Caenorhabditis elegans*.



19.9 The Principle of Cytoplasmic Segregation The unequal distribution of some component in the cytoplasm of a cell may determine the fates of its descendants.

LENS DIFFERENTIATION IN THE VERTEBRATE EYE The development of the lens in the vertebrate eye is a classic example of induction. In a frog embryo, the developing forebrain bulges out at both sides to form the *optic vesicles*, which expand until they come into contact with the cells at the surface of the head (Figure 19.10). The surface tissue in the region of contact thickens, forming a *lens placode*—tissue that will ultimately form the lens. The lens placode bends inward, folds over on itself, and ultimately detaches from the surface tissue to produce a structure that will develop into the lens. If the growing optic vesicle is cut away before it contacts the surface cells, no lens forms. Placing an impermeable barrier between the optic vesicle and the surface cells also prevents the lens from forming. These observations suggest that the surface tissue begins to develop into a lens when it receives a signal from the optic vesicle. Such a signal is termed an **inducer**.



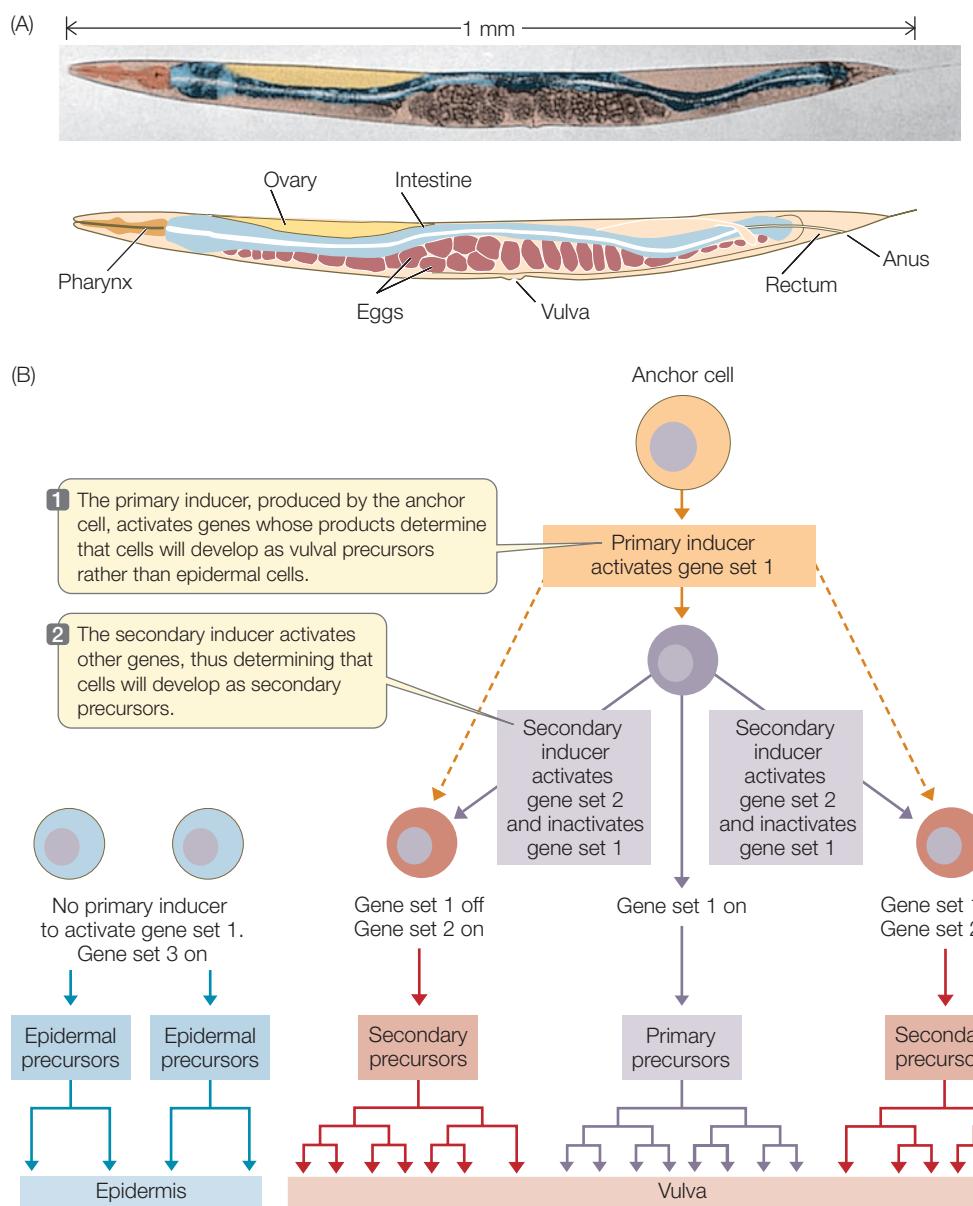
19.10 Embryonic Inducers in Vertebrate Eye Development The eye of a frog develops as different cells induce changes in neighboring cells.

Inducers trigger sequences of gene expression in the responding cells. How cells switch on different sets of genes that govern development and direct the formation of body plans is of great interest to developmental and evolutionary biologists. They use model organisms to investigate the major principles governing these processes.

VULVAL DIFFERENTIATION IN THE NEMATODE The tiny nematode *Caenorhabditis elegans* is a favorite model organism for studying development. Its genome was one of the first eukaryotic genomes to be sequenced (see Section 17.3). It develops from fertilized egg to larva in only about 8 hours, and the worm reaches the adult stage in just 3.5 days. The process is easily observed using a low-magnification dissecting microscope because the body covering is transparent (Figure 19.11A).

The adult nematode is *hermaphroditic*, containing both male and female reproductive organs. It lays eggs through a pore called the *vulva* on the ventral (lower) surface. During development, a single cell, called the *anchor cell*, induces the vulva to form from six cells on the worm's ventral surface. In this case, there are two molecular signals, the *primary inducer* and the *secondary inducer* (or *lateral signal*). Each of the six ventral cells has three possible fates: it may become a primary vulval precursor cell, a secondary vulval precursor cell, or simply become part of the worm's skin—an epidermal cell. You can follow the sequence of events in Figure 19.11B. The concentration gradient of the primary inducer, LIN-3, is key: the anchor cell produces LIN-3, which diffuses out of the cell and forms a concentration gradient with respect to adjacent cells. Cells that receive the most LIN-3 become vulval precursor cells; cells slightly farther from the anchor cell receive less LIN-3 and become epidermal cells. Induction involves the activation or inactivation of specific genes through signal transduction cascades in the responding cells (Figure 19.12).

Nematode development illustrates the important observation that *much of development is controlled by molecular switches that allow a cell to proceed down one of two alternative tracks*. One challenge for developmental biologists is to find these switches and determine how they work. The primary inducer, LIN-3, released by the *C. elegans* anchor cell is a growth factor homologous to a ver-



19.11 Induction during Vulval Development in *Caenorhabditis elegans*

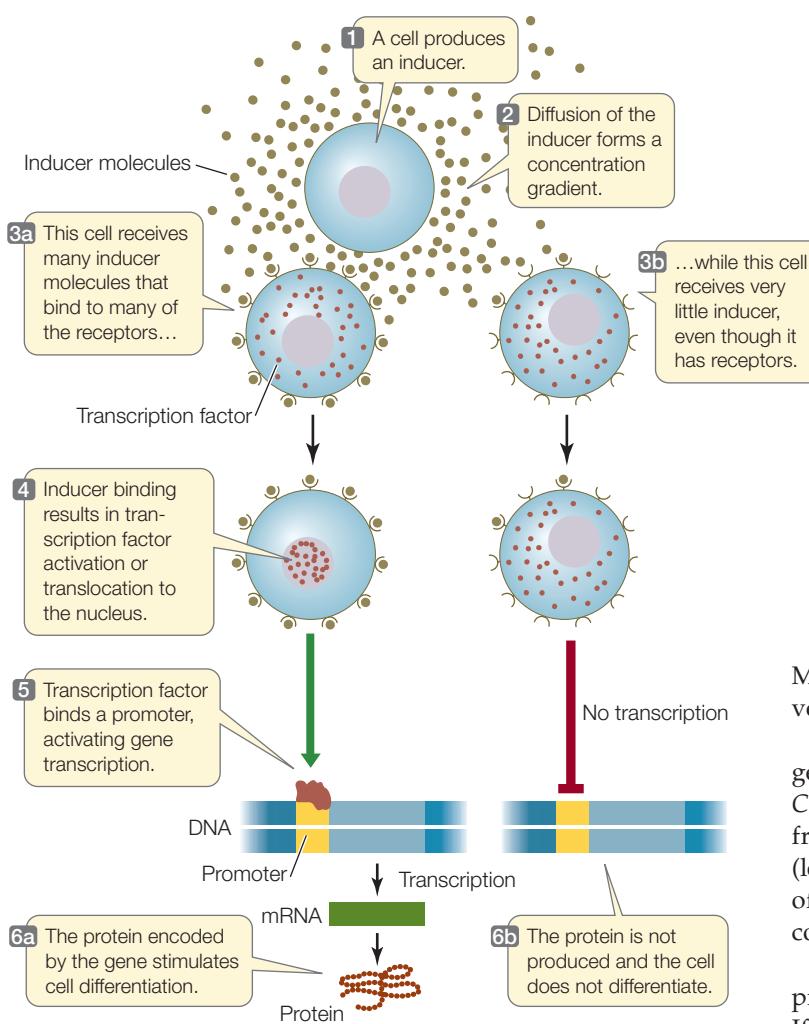
elegans (A) In the nematode *C. elegans*, it has been possible to follow all of the cell divisions from the fertilized egg to the 959 cells found in the fully developed adult. (B) During vulval development, a molecule secreted by the anchor cell (the LIN-3 protein) acts as the primary inducer. The primary precursor cell (the one that received the highest concentration of LIN-3) then secretes a secondary inducer (the lateral signal) that acts on its neighbors. The gene expression patterns triggered by these molecular switches determine cell fates.

tebrate growth factor called EGF (epidermal growth factor). LIN-3 binds to a receptor on the surfaces of vulval precursor cells, setting in motion a signal transduction cascade involving the Ras protein and MAP kinases (see Figure 7.12). This results in increased transcription of the genes involved in the differentiation of vulval cells.

19.4 RECAP

Cellular differentiation involves cytoplasmic segregation and induction. Cytoplasmic segregation is the unequal distribution of gene products in the egg, zygote, or early embryo. Induction occurs when one cell or tissue sends a chemical signal to another.

- How does cytoplasmic segregation result in polarity in a fertilized egg, and how does polarity affect cell differentiation? See pp. 413–414 and Figure 19.9
- Describe an example of how induction influences tissue formation in the vertebrate eye. See p. 416 and Figure 19.10
- How do inducer molecules interact with transcription factors to produce differentiated cells? See p. 416 and Figure 19.12



19.12 Induction The concentration of an inducer directly affects the degree to which a transcription factor is activated. The inducer acts by binding to a receptor on the target cell. This binding is followed by signal transduction involving transcription factor activation or translocation from the cytoplasm to the nucleus. In the nucleus it acts to stimulate the expression of genes involved in cell differentiation.

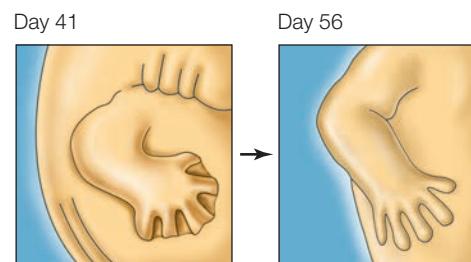
We have seen that cytoplasmic segregation and induction lead to cell differentiation, and have seen two examples of how these processes lead to organ formation in developing multicellular organisms. We now take a closer look at how gene expression affects differentiation and development.

19.5 How Does Gene Expression Determine Pattern Formation?

Pattern formation is the process that results in the spatial organization of a tissue or organism. It is inextricably linked to morphogenesis, the creation of body form. You might expect morphogenesis to involve a lot of cell division, followed by differentiation—and it does. But what you might not expect is the amount of programmed cell death—apoptosis—that occurs during morphogenesis.

Multiple genes interact to determine developmental programmed cell death

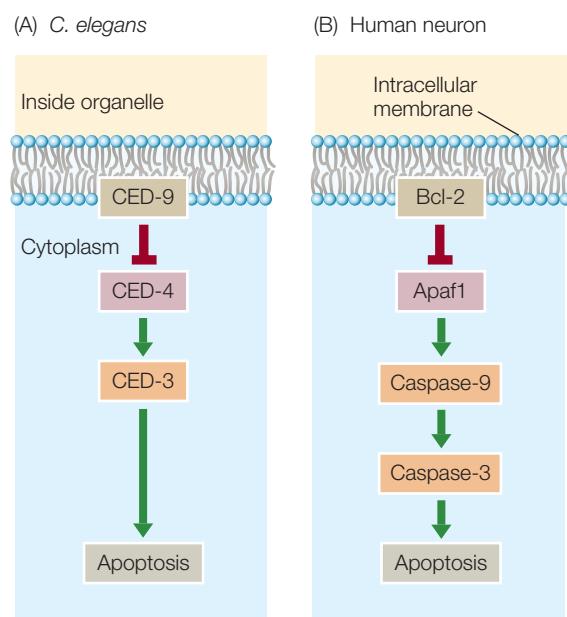
We noted in Section 11.6 that apoptosis is a programmed series of events that leads to cell death. Apoptosis is an integral part of the normal development and life of an organism. For example, in an early human embryo, the hands and feet look like tiny paddles: the tissues that will become fingers and toes are linked by connective tissue. Between days 41 and 56 of development, the cells between the digits die, freeing the individual fingers and toes:



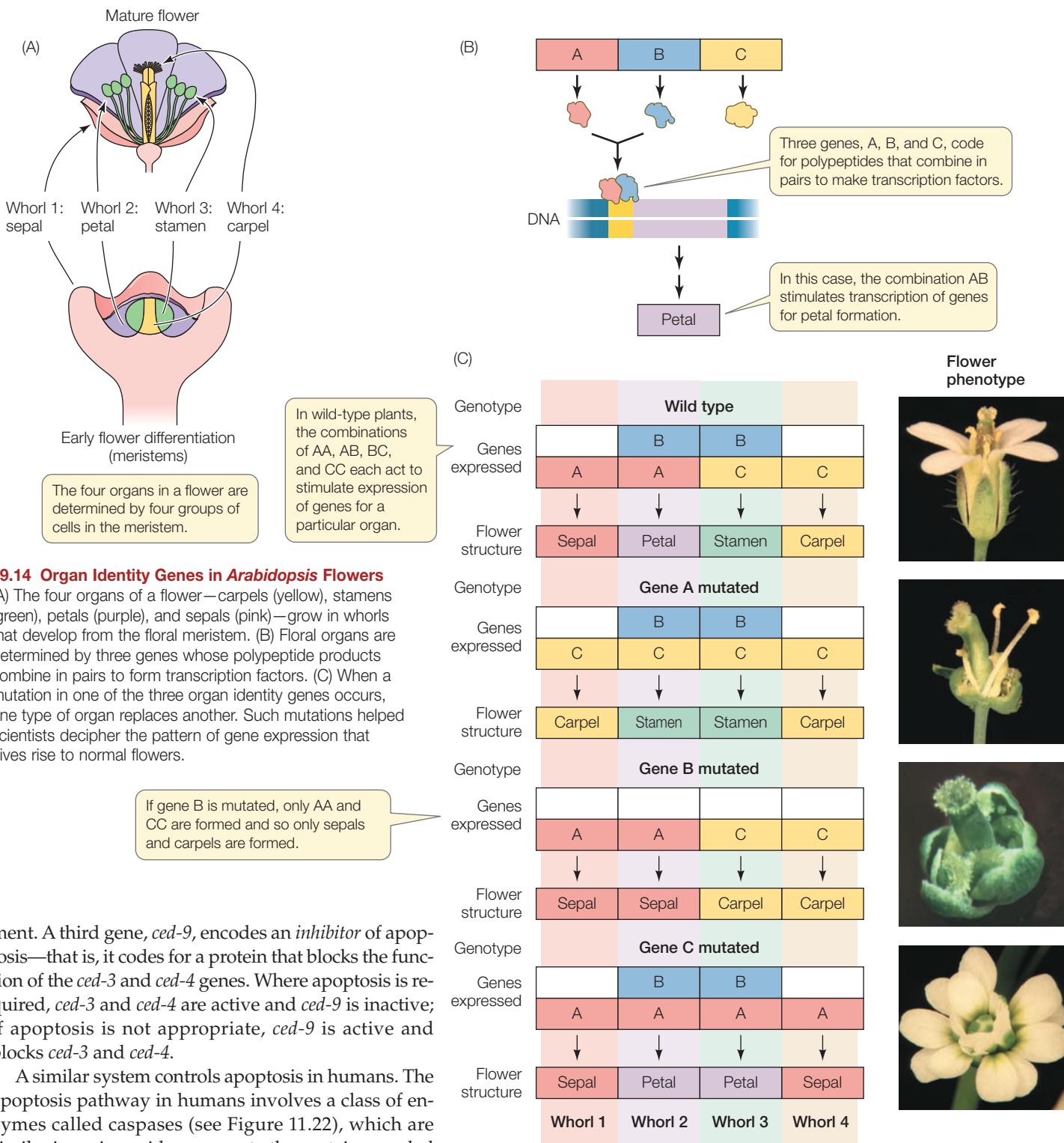
Many cells and structures form and then disappear during development, in processes involving apoptosis.

Model organisms have been very useful in studying the genes involved in apoptosis. For example, the nematode worm *C. elegans* produces precisely 1,090 somatic cells as it develops from a fertilized egg into an adult, but 131 of those cells die (leaving 959 cells in the adult worm). The sequential expression of two genes called *ced-4* and *ced-3* (for cell death) appears to control this programmed cell death (**Figure 19.13**).

In the nematode nervous system, 302 neurons come from 405 precursors; thus 103 neural precursor cells undergo apoptosis. If the protein encoded by either *ced-3* or *ced-4* is nonfunctional, all 405 cells form neurons, resulting in abnormal brain develop-



19.13 Pathways for Apoptosis In the worm *C. elegans* (A) and humans (B) similar pathways for apoptosis are controlled by genes with similar sequences and functions.



19.14 Organ Identity Genes in *Arabidopsis* Flowers

(A) The four organs of a flower—carpels (yellow), stamens (green), petals (pink), and sepals (pink)—grow in whorls that develop from the floral meristem. (B) Floral organs are determined by three genes whose polypeptide products combine in pairs to form transcription factors. (C) When a mutation in one of the three organ identity genes occurs, one type of organ replaces another. Such mutations helped scientists decipher the pattern of gene expression that gives rise to normal flowers.

ment. A third gene, *ced-9*, encodes an *inhibitor* of apoptosis—that is, it codes for a protein that blocks the function of the *ced-3* and *ced-4* genes. Where apoptosis is required, *ced-3* and *ced-4* are active and *ced-9* is inactive; if apoptosis is not appropriate, *ced-9* is active and blocks *ced-3* and *ced-4*.

A similar system controls apoptosis in humans. The apoptosis pathway in humans involves a class of enzymes called caspases (see Figure 11.22), which are similar in amino acid sequence to the protein encoded by *ced-3* in *C. elegans*. Humans have one protein, Bcl-2, that inhibits apoptosis and is similar to the product of *ced-9*, and another protein, Apaf1, that stimulates apoptosis like CED-4. As the human nervous system develops, half of the neurons that are formed undergo apoptosis. So humans and nematodes, two species separated by more than 600 million years of evolutionary history, have similar genes controlling programmed cell death (see Figure 19.13). The commonality of this pathway indicates its importance: mutations are harmful and evolution selects against them.

Plants have organ identity genes

Like animals, plants have organs—for example, leaves and roots. Many plants form flowers, and many flowers are composed of four types of organs: sepals, petals, stamens (male reproductive organs), and carpels (female reproductive organs). These floral organs occur in concentric *whorls*, with groups of each organ type encircling a central axis. The sepals are on the outside and the carpels are on the inside (Figure 19.14A).

In the model plant *Arabidopsis thaliana* (thale cress), flowers develop in a radial pattern around the shoot apex as it develops and elongates. The whorls develop from a *meristem* of about 700 undifferentiated cells arranged in a dome, which is at the growing point on the stem. How is the identity of a particular whorl determined? A group of genes called **organ identity genes** encode proteins that act in combination to produce specific whorl features (Figure 19.14B and C):

- Genes in class A are expressed in whorls 1 and 2 (which form sepals and petals, respectively).
- Genes in class B are expressed in whorls 2 and 3 (which form petals and stamens).
- Genes in class C are expressed in whorls 3 and 4 (which form stamens and carpels).

Two lines of experimental evidence support this model of organ identity gene function:

- *Loss-of-function mutations*: for example, a mutation in a class A gene results in no sepals or petals. In any organism, the replacement of one organ for another is called *homeosis*, and this type of mutation is a **homeotic mutation** (see Figure 19.14C).
- *Gain-of-function mutations*: for example, a promoter for a class C gene can be artificially coupled to a class A gene. In this case, the class A gene is expressed in all four whorls, resulting in only sepals and petals.

Genes in classes A, B, and C code for transcription factors that are active as dimers, that is, proteins with two polypeptide subunits. Gene regulation in these cases is *combinatorial*—that is, the composition of the dimer determines which genes will be activated. For example, a dimer made up of two class A monomers activates transcription of the genes that make sepals; a dimer made up of A and B monomers results in petals, and so forth. A common feature of the A, B, and C proteins, as well as many other plant transcription factors, is a DNA-binding domain called the **MADS box**. These proteins also have domains that can bind to other proteins in a *transcription initiation complex*. As we discuss in Chapter 16, transcription initiation in eukaryotes is controlled by a complex of proteins that interact with DNA and other proteins at the promoter. The MADS box proteins participate in this complex to control the expression of specific genes.

Some familiar ornamental plants have mutations in floral organ identity genes. For example, many rose varieties have mutations in a C gene, resulting in multiple rows of petals instead of the single set of five petals found in wild roses. An understanding of the molecular basis of floral organ identity may have practical uses. Many of the foods that make up the human diet come from fruits and seeds, which form from parts of the carpel—the female reproductive organ of the flower. Genetically modifying plants to produce more carpels could increase the amount of fruit or grain a crop produces. A genetic system similar to the one described here for *Arabidopsis* controls floral organ formation in rice, humanity's most widely consumed plant. Appropriate mutations in these genes might lead to more grain produced per plant.

Transcription of the floral organ identity genes is controlled by other gene products, including the LEAFY protein. Plants with loss-of-function mutations in the LEAFY gene make flow-

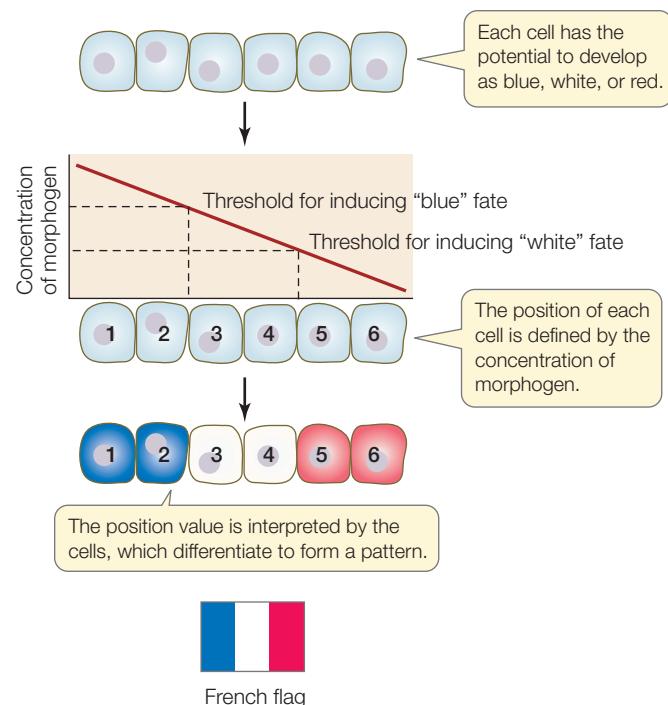
ering stems instead of flowers, with increased numbers of modified leaves called bracts. The wild-type LEAFY protein acts as a transcription factor, stimulating expression of the class A, B, and C genes so that they produce flowers. This finding, too, has practical applications. It usually takes 6–20 years for a citrus tree to produce flowers and fruits. Scientists have made transgenic orange trees expressing the LEAFY gene coupled to a strongly expressed promoter. These trees flower and fruit years earlier than normal trees.

Morphogen gradients provide positional information

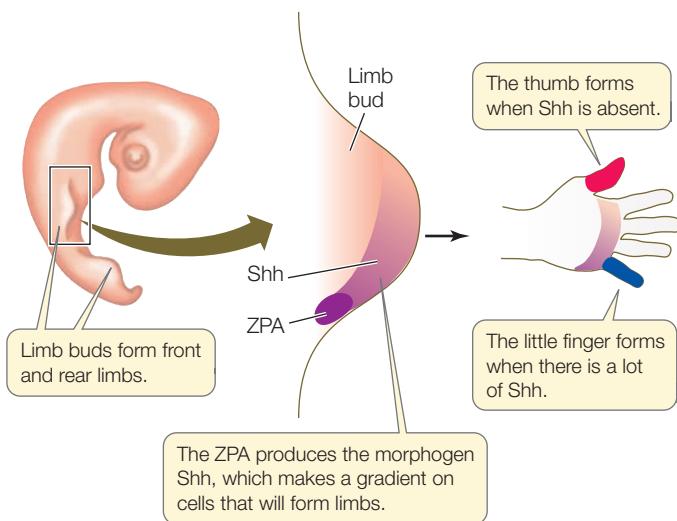
During development, the key cellular question, “What am I (or what will I be)?” is often answered in part by “Where am I?” Think of the cells in the developing nematode, which develop into different parts of the vulva depending on their positions relative to the anchor cell. This spatial “sense” is called **positional information**. Positional information often comes in the form of an inducer called a **morphogen**, which diffuses from one group of cells to surrounding cells, setting up a concentration gradient. There are two requirements for a signal to be considered a morphogen:

- It must directly affect target cells, rather than triggering a secondary signal that affects target cells.
- Different concentrations of the signal must cause different effects.

Developmental biologist Lewis Wolpert uses the “French flag model” to explain morphogens (Figure 19.15). This model can be applied to the differentiation of the vulva in *C. elegans* (see Figure 19.11) and to the development of vertebrate limbs.



19.15 The French Flag Model In the “French flag” model, a concentration gradient of a diffusible morphogen signals each cell to specify its position.



The vertebrate limb develops from a paddle-shaped *limb bud* (Figure 19.16). The cells that develop into different digits must receive positional information; if they do not, the limb will be totally disorganized (imagine a hand with only thumbs or only little fingers). A group of cells at the posterior base of the limb bud, just where it joins the body wall, is called the *zone of polarizing activity* (ZPA). The cells of the ZPA secrete a morphogen called *Sonic hedgehog* (Shh), which forms a gradient that determines the posterior–anterior (little finger to thumb) axis of the developing limb. The cells getting the highest dose of Shh form the little finger; those getting the lowest dose develop into the thumb. Recall the French flag model when considering the gradient of Shh.

A cascade of transcription factors establishes body segmentation in the fruit fly

Perhaps the best-studied example of how gene expression affects cell fate in response to morphogens is body segmentation in the fruit fly *Drosophila melanogaster*. The body segments of this model organism are clearly different from one another. The adult fly has an anterior *head* (composed of several fused segments), three different *thoracic* segments, and eight *abdominal* segments at the posterior end. Each segment develops into different body parts: for example, antennae and eyes develop from head segments, wings from the thorax, and so on.

The life cycle of *Drosophila* from fertilized egg to adult takes about 2 weeks at room temperature. The egg hatches into a larva, which then forms a pupa, which finally is transformed into the adult fly. By the time a larva appears—about 24 hours after fertilization—there are recognizable segments. The thoracic and abdominal segments all look similar, but *the fates of the cells to become different adult segments is already determined*. The determination events in the first 24 hours will be our focus here.

Several types of genes are expressed sequentially in the embryo to define these segments:

19.16 Specification of the Vertebrate Limb and the French Flag

Model The zone of polarizing activity (ZPA) in the limb bud of the embryo secretes the morphogen Sonic hedgehog (Shh). Cells in the bud form different digits depending on the concentration of Shh.

- First, cells in the mother that are adjacent to the maturing egg release products that set up anterior–posterior and dorsal–ventral axes in the egg.
- Next, a series of gene products in the embryo successively define the position of each cell in a segment relative to these axes. For example, a cell might first be defined as being in the head rather than in the abdomen in the anterior–posterior axis; then it might be defined as being on the ventral (top) side of the head.
- Finally, a set of genes called *Hox genes* control the ultimate identity of each body part; for example, determining that the cells at a particular position in the head will make mouthparts.

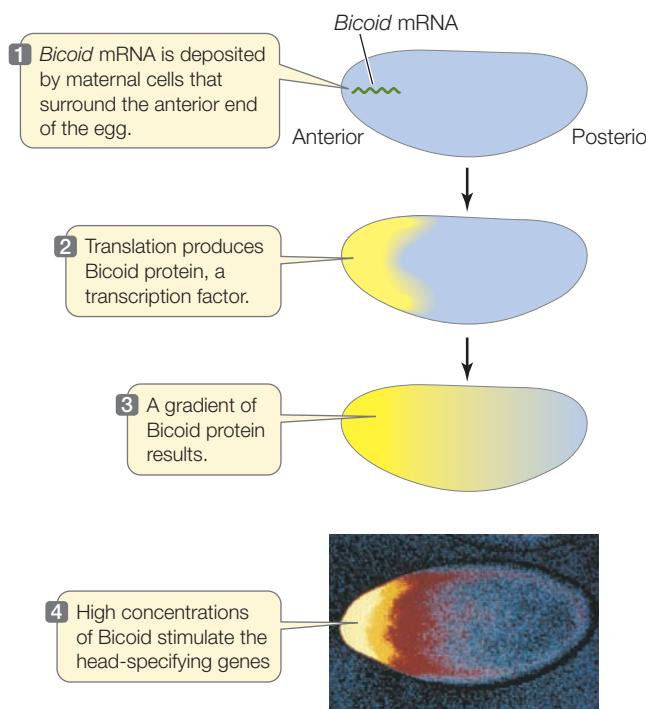
The genes involved in each of these steps code for transcription factors, which in turn control the synthesis of other transcription factors acting on the next set of genes. This cascade of events may remind you of a signal transduction cascade (see Section 7.3), only in this case it is a *cascade of events that occurs over time and location*, rather than abruptly and in a single cell. The genes finally expressed are the ones familiar to you: they code for protein kinases, receptors, and other proteins that carry out the functions of the cell.

The description of these events in fruit fly development is one of the great achievements in modern biology. It gave biologists a deep understanding of how the events that specify cell identity unfold. We will only skim the surface of the process here, but keep in mind the basic principle of a transcriptional cascade. As we will see in Chapter 20, the fruit fly has been a true model organism in this case, because these findings have informed research on other organisms, including mammals.

Experimental genetics was used to elucidate the events leading to cell fate determination in *Drosophila*:

- First, developmental mutations were identified. For example, a mutant strain might produce larvae with two heads or no segments.
- Then, the mutant was compared with wild-type flies, and the gene responsible for the developmental mistake, and its protein product (if appropriate), were isolated.
- Finally, experiments with the gene (making transgenic flies) and protein (injecting the protein into an egg or into an embryo) were done to confirm the proposed developmental pathway.

Together, these approaches revealed a sequential pattern of gene expression that results in the determination of each segment within 24 hours after fertilization. Several classes of genes are involved.



MATERNAL EFFECT GENES Like the eggs and early embryos of sea urchins, *Drosophila* eggs and larvae are characterized by unevenly distributed cytoplasmic determinants (see Figure 19.9). These molecular determinants, which include both mRNAs and proteins, are the products of specific **maternal effect genes**. These genes are transcribed in the cells of the mother's ovary that surround what will be the anterior portion of the egg. The transcription products are passed to the egg by cytoplasmic bridges. Two maternal effect genes, called *Bicoid* and *Nanos*, help determine the anterior-posterior axis of the egg. (The dorsal-ventral [back-belly] axis is determined by other maternal effect genes that will not be described here.)

The mRNAs for *Bicoid* and *Nanos* diffuse from the mother's cells into what will be the anterior end of the egg. The *Bicoid* mRNA is translated to produce Bicoid protein, which diffuses away from the anterior end, establishing a gradient in the egg cytoplasm (Figure 19.17). Where it is present in sufficient concentration, Bicoid acts as a transcription factor to stimulate the transcription of the *Hunchback* gene in the early embryo. A gradient of the Hunchback protein establishes the head, or anterior, region.

Meanwhile, the egg's cytoskeleton transports the *Nanos* mRNA from the anterior end of the egg, where it was deposited, to the posterior end, where it is translated. The Nanos protein forms a gradient with the highest concentration at the posterior end. At that end, the Nanos protein inhibits the translation of *Hunchback* mRNA. Thus, the action of both Bicoid and Nanos establish the Hunchback gradient, which determines the anterior and posterior ends of the embryo.

19.17 Bicoid Protein Provides Positional Information The anterior-posterior axis of *Drosophila* arises from the gradient of a morphogen encoded by *Bicoid*, a maternal effect gene. Bicoid protein is also a transcription factor, which activates a gene to specify that the anterior region will become the head of the fly. Other maternal effect genes in the posterior region of the embryo inhibit Bicoid, thus limiting its activity in that region.

How did biologists elucidate these pathways? Let's look more closely at the experimental approaches used in this case.

- Females that are homozygous for a particular *bicoid* mutation produce larvae with no head and no thorax; thus the Bicoid protein must be needed for the anterior structures to develop.
- If the eggs of these *bicoid* mutants are injected at the anterior end with cytoplasm from the anterior region of a wild-type egg, the injected eggs develop into normal larvae. This experiment also shows that the Bicoid protein is involved in the development of anterior structures.
- If cytoplasm from the anterior region of a wild-type egg is injected into the posterior region of another egg, anterior structures develop there. The degree of induction depends on how much cytoplasm is injected.
- Eggs from homozygous *nanos* mutant females develop into larvae with missing abdominal segments.
- If cytoplasm from the posterior region of a wild-type egg is injected into the posterior region of a *nanos* mutant egg, it will develop normally.

These and other experiments led scientists to understand the cascade of events that determine cell fates.

The events involving *Bicoid*, *Nanos*, and *Hunchback* begin before fertilization and continue after it, during the multinucleate stage, which lasts a few hours. At this stage, the embryo looks like a bunch of indistinguishable nuclei under the light microscope. But the cell fates have already begun to be determined. After the anterior and posterior ends have been established, the next step in pattern formation is the determination of segment number and locations.

SEGMENTATION GENES The number, boundaries, and polarity of the *Drosophila* larval segments are determined by proteins encoded by the **segmentation genes**. These genes are expressed when there are about 6,000 nuclei in the embryo (about three hours after fertilization). Three classes of segmentation genes act one after the other to regulate finer and finer details of the segmentation pattern:

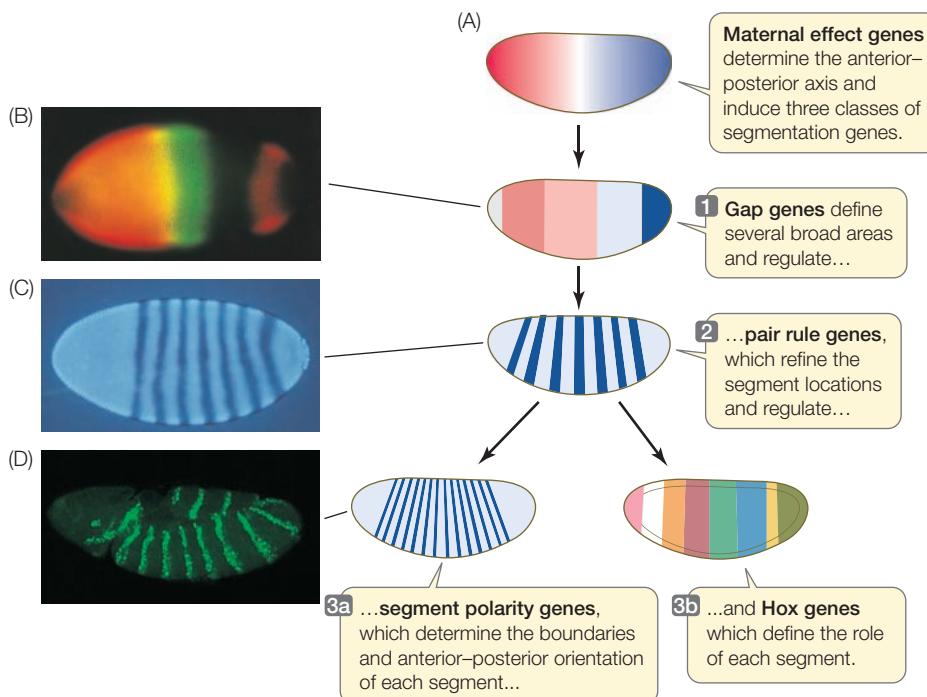
- **Gap genes** organize broad areas along the anterior-posterior axis. Mutations in gap genes result in gaps in the body plan—the omission of several consecutive larval segments.
- **Pair rule genes** divide the embryo into units of two segments each. Mutations in pair rule genes result in embryos missing every other segment.

19.18 A Gene Cascade Controls Pattern Formation in the *Drosophila* Embryo

(A) Maternal effect genes (see Figure 19.17) induce gap, pair rule, and segment polarity genes—collectively referred to as segmentation genes. (B) Two gap genes, *Hunchback* (orange) and *Krüppel* (green) overlap; both genes are transcribed in the yellow area. (C) The pair rule gene *Fushi tarazu* is transcribed in the dark blue areas. (D) The segment polarity gene *Engrailed* (bright green) is seen here at a slightly more advanced stage than is depicted in (A). By the end of this cascade, a group of nuclei at the anterior of the embryo, for example, is determined to become the first head segment in the adult fly.

yourBioPortal.com

GO TO Animated Tutorial 19.3 •
Pattern Formation in the *Drosophila* Embryo

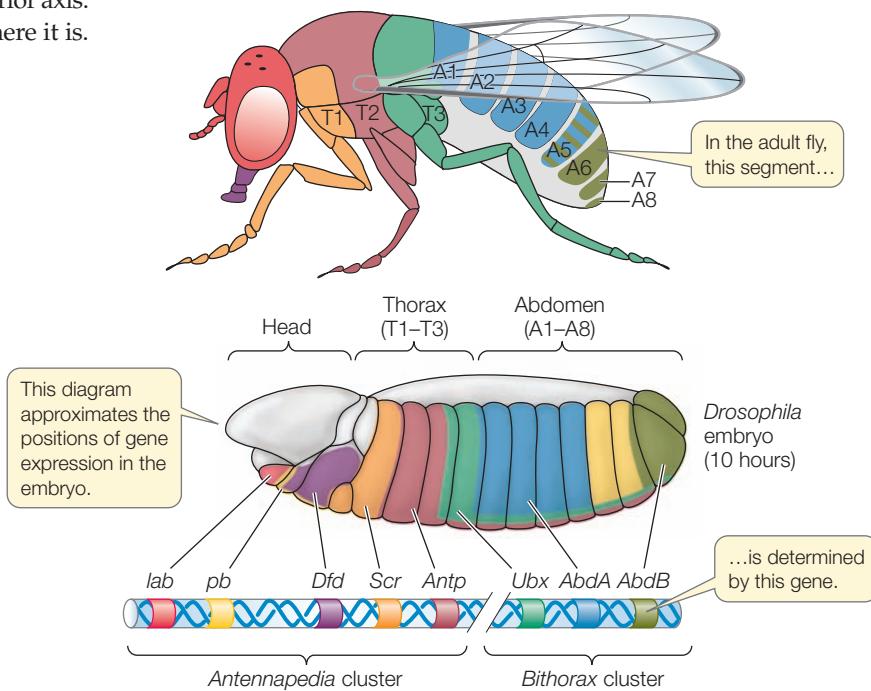


- **Segment polarity genes** determine the boundaries and anterior–posterior organization of the individual segments. Mutations in segment polarity genes can result in segments in which posterior structures are replaced by reversed (mirror-image) anterior structures.

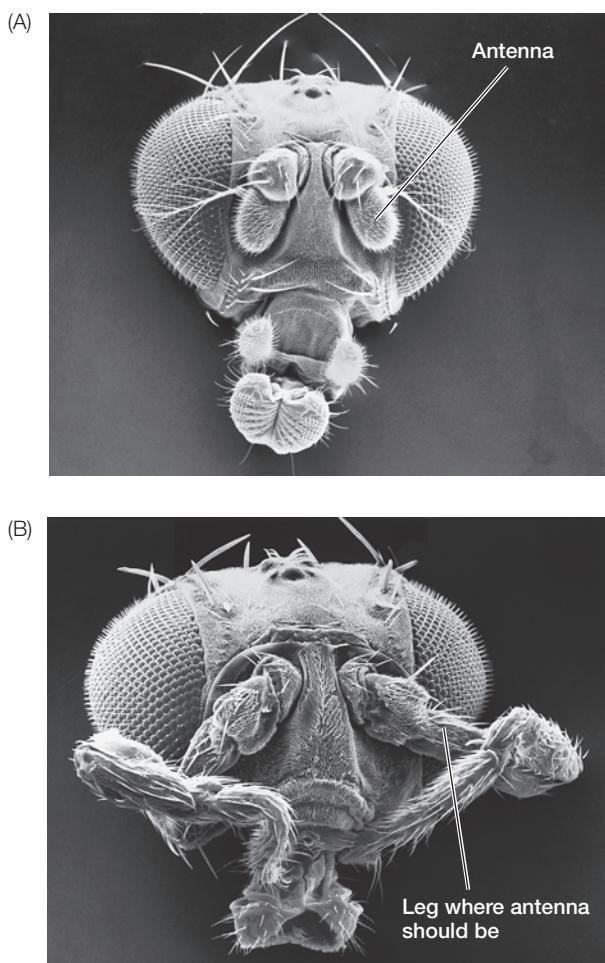
The expression of these genes is sequential (**Figure 19.18**). The maternal effect protein Bicoid, which begins the cascade, acts as a morphogen and transcription factor to stimulate the expression of genes such as *Hunchback* that set up the anterior–posterior axis. As a result, a nucleus in the early embryo “knows” where it is. The *Hunchback* protein stimulates gap gene transcription, the products of the gap genes activate pair rule genes, and the pair rule gene products activate segment polarity genes. By the end of this cascade, nuclei throughout the embryo “know” which segment they will be part of in the adult fly.

The next set of genes in the cascade determines the form and function of each segment.

HOX GENES **Hox genes** encode a family of transcription factors that are expressed in different combinations along the length of the embryo, and help determine cell fate within each segment. Hox gene expression tells the cells of a segment in the head to make eyes, those of a segment in the thorax to make wings, and so on. The *Drosophila* Hox genes occur in two clusters on chromosome 3, in the same order as the segments whose function they determine (**Figure 19.19**). By the time the fruit fly larva hatches, its segments are completely determined. Hox genes are homeotic genes that are shared by all animals, and they are functionally analogous to the organ identity genes of plants. However, they differ from plant homeotic genes in DNA sequence and encoded



19.19 Hox Genes in *Drosophila* Determine Segment Identity Two clusters of Hox genes on chromosome 3 (center) determine segment function in the adult fly (top). These genes are expressed in the embryo (bottom) long before the structures of the segments actually appear.



protein structure. This is not surprising, given that the last common ancestor of plants and animals was unicellular, and therefore multicellularity evolved independently for plants and animals.

In *Drosophila*, the maternal effect genes, segmentation genes, and Hox genes interact to “build” a larva step by step, beginning with the unfertilized egg. How do we know that the Hox genes determine segment identity? An important clue came from bizarre homeotic mutations observed in *Drosophila*. The *antennapedia* mutation causes legs to grow on the head in place of antennae (Figure 19.20), and the *bithorax* mutation causes an extra pair of wings to grow in a thoracic segment where wings do not normally occur (see Figure 20.3). Edward Lewis at Caltech found that *antennapedia* and *bithorax* mutations resulted from changes in Hox genes.

The first cluster of Hox genes—the *Antennapedia* cluster—specifies anterior segments, starting with genes for the different head segments and ending with thoracic segments. The second cluster (*Bithorax*) contains three genes. It begins with a gene specifying the last thoracic segment, followed by a gene for the anterior abdominal segments, and ends with a gene for the posterior abdominal segments (see Figure 19.19). Lewis hypothe-

19.20 A Homeotic Mutation in *Drosophila* Mutations of the Hox genes cause body parts to form on inappropriate segments. (A) A wild-type fruit fly. (B) An *antennapedia* mutant fruit fly. Mutations such as this reveal the normal role of the *Antennapedia* gene in determining segment function.

thesized that all of the Hox genes might have come from the duplication of a single gene in an ancestral, unsegmented organism. Since Lewis put forward this hypothesis, molecular research methods became available to test it.

Hox genes encode transcription factors

Molecular biologists confirmed Lewis’s hypothesis using nucleic acid hybridization. Several scientists found that probes for a sequence found in the *Bithorax* cluster could bind to other sequences in both the *Bithorax* and *Antennapedia* clusters. In other words, this DNA sequence is common to all the Hox genes in both clusters. It is also found in several of the segmentation genes, as well as other genes that encode transcription factors.

This 180-base-pair DNA sequence is called the **homeobox**. It encodes a 60-amino acid sequence called the *homeodomain*. The homeodomain recognizes and binds to a specific DNA sequence in the promoters of its target genes. However, this recognition is usually not sufficient to allow the transcription factor to bind fully to a promoter and turn the target gene on or off. Other transcription factors are also involved.

The Hox genes are found in animals with an anterior–posterior axis, where they play a role in development similar to that played by MADS box genes in plants. But the homeobox is found in many different transcription factors, including some from plants. The evolutionary significance of these common pathways for development will be discussed in the next chapter.

19.5 RECAP

A cascade of transcription factors governs pattern formation and the subsequent development of animal and plant organs. Often these transcription factors create or respond to morphogen gradients. In plants, cell fate is often determined by MADS box genes, and in animal embryos, cell fate is determined in part by Hox genes.

- How is apoptosis crucial in shaping the developing embryo? **See p. 417**
- How do organ identity genes act in *Arabidopsis*? **See pp. 418–419 and Figure 19.14**
- List the key attributes of a morphogen. How does the Bicoid protein fit this definition? **See pp. 420–421 and Figure 19.17**
- How is segment identity established in the *Drosophila* embryo? **Review pp. 421–422 and Figure 19.19**

CHAPTER SUMMARY

19.1 What Are the Processes of Development?

- A multicellular organism begins its development as an embryo. A series of embryonic stages precedes the birth of an independent organism. **Review Figure 19.1, WEB ACTIVITY 19.1**
- The processes of development are **determination, differentiation, morphogenesis, and growth**.
- **Differential gene expression** is responsible for the differences between cell types. **Cell fate** is determined by environmental factors, such as the cell's position in the embryo, as well as by intracellular influences. **Review Figure 19.2**
- Determination is followed by differentiation, the actual changes in biochemistry, structure, and function that result in cells of different types. Determination is a commitment; differentiation is the realization of that commitment.

19.2 Is Cell Differentiation Irreversible?

- The zygote is **totipotent**; it is capable of forming every type of cell in the adult body.
- The ability to create **clones** from differentiated cells demonstrates the principle of **genomic equivalence**. **Review Figures 19.3 and 19.4**
- **Stem cells** produce daughter cells that differentiate when provided with appropriate intercellular signals. Some **multipotent** stem cells in the adult body can differentiate into a limited number of cell types to replace dead cells and maintain tissues. **Review Figure 19.5**
- Embryonic stem cells are **pluripotent** and can be cultured in the laboratory. Under suitable environmental conditions, these cells can differentiate into any tissue type. **Induced pluripotent stem cells** have similar characteristics. This has led to technologies to replace cells or tissues damaged by injury or disease. **Review Figure 19.6, ANIMATED TUTORIAL 19.1**

19.3 What Is the Role of Gene Expression in Cell Differentiation?

- Differential gene expression results in cell differentiation. Transcription factors are especially important in regulating gene expression during differentiation.
- Complex interactions of many genes and their products are responsible for differentiation during development. **Review Figure 19.7**

19.4 How Is Cell Fate Determined?

- **Cytoplasmic segregation**—the unequal distribution of **cytoplasmic determinants** in the egg, zygote, or early embryo—can establish **polarity** and lead to cell fate determination. **Review Figures 19.8 and 19.9, ANIMATED TUTORIAL 19.2**
- **Induction** is a process by which embryonic animal tissues direct the development of neighboring cells and tissues by secreting chemical signals, called **inducers**. **Review Figure 19.10**
- The induction of the vulva in the nematode *Caenorhabditis elegans* offers an example of how inducers act as molecular switches to direct a cell down one of two differentiation paths. **Review Figures 19.11 and 19.12**

19.5 How Does Gene Expression Determine Pattern Formation?

- **Pattern formation** is the process that results in the spatial organization of a tissue or organism.
- During development, selective elimination of cells by apoptosis results from the expression of specific genes. **Review Figure 19.13**
- Sepals, petals, stamens, and carpels form in plants as a result of combinatorial interactions between transcription factors encoded by **organ identity genes**. **Review Figure 19.14**
- The transcription factors encoded by floral organ identity genes contain an amino acid sequence called the **MADS box** that can bind to DNA.
- Both plants and animals use **positional information** as a basis for pattern formation. Positional information usually comes in the form of a signal called a **morphogen**. Different concentrations of the morphogen cause different effects. **See Figures 19.15 and 19.16**
- In the fruit fly *D. melanogaster*, a cascade of transcriptional activation sets up the axes of the embryo, the development of the segments, and finally the determination of cell fate in each segment. The cascade involves the sequential expression of maternal effect genes, gap genes, pair rule genes, segment polarity genes, and Hox genes. **Review Figures 19.18 and 19.19, ANIMATED TUTORIAL 19.3**
- Hox genes help to determine cell fate in the embryos of all animals. The **homeobox** is a DNA sequence found in Hox genes and other genes that code for transcription factors. The sequence of amino acids encoded by the homeobox is called the **homeodomain**.

SELF-QUIZ

1. Which statement about determination is true?
 - Differentiation precedes determination.
 - All cells are determined after two cell divisions in most organisms.
 - A determined cell will keep its determination no matter where it is placed in an embryo.
 - A cell changes its appearance when it becomes determined.
 - A differentiated cell has the same pattern of transcription as a determined cell.
2. Cloning experiments on sheep, frogs, and mice have shown that
 - nuclei of adult cells are totipotent.
 - nuclei of embryonic cells can be totipotent.
 - nuclei of differentiated cells have different genes than zygote nuclei have.
 - differentiation is fully reversible in all cells of a frog.
 - differentiation involves permanent changes in the genome.

3. The term “induction” describes a process in which a cell or cells
 - a. influence the development of another group of cells.
 - b. trigger cell movements in an embryo.
 - c. stimulate the transcription of their own genes.
 - d. organize the egg cytoplasm before fertilization.
 - e. inhibit the movement of the embryo.
4. Stem cells from adult animals
 - a. are always totipotent.
 - b. divide when provided with external signals.
 - c. are not present in bone marrow.
 - d. are present in an embryo but not an adult.
 - e. can be turned into differentiated cells with only a few genes.
5. Which statement about cytoplasmic determinants in *Drosophila* is *not* true?
 - a. They specify the dorsal–ventral and anterior–posterior axes of the embryo.
 - b. Their positions in the embryo are determined by cytoskeletal action.
 - c. Some are products of specific genes in the mother fruit fly.
 - d. They do not produce gradients.
 - e. They have been studied by the transfer of cytoplasm from egg to egg.
6. In fruit flies, the following genes are used to determine segment polarity: (k) gap genes; (l) Hox genes; (m) maternal effect genes; (n) pair rule genes. In what order are these genes expressed during development?
 - a. k, l, m, n
 - b. l, k, n, m
 - c. m, k, n, l
 - d. n, k, m, l
 - e. n, m, k, l
7. Which statement about induction is *not* true?
 - a. One group of cells induces adjacent cells to develop in a certain way.
 - b. It triggers a sequence of gene expression in target cells.
 - c. Single cells cannot form an inducer.
 - d. A tissue may be induced as well as make an inducer.
 - e. The chemical identification of specific inducers has not been achieved.
8. In the process of pattern formation in the *Drosophila* embryo,
 - a. the first steps are specified by Hox genes.
 - b. mutations in pair rule genes result in embryos missing every other segment.
 - c. mutations in gap genes result in the insertion of extra segments.
 - d. segment polarity genes determine the dorsal–ventral axes of segments.
 - e. all segments develop the same organs.
9. Homeotic mutations
 - a. are often severe and result in structures at inappropriate places.
 - b. cause subtle changes in the forms of larvae or adults.
 - c. occur only in prokaryotes.
 - d. do not affect the animal’s DNA.
 - e. are confined to the zone of polarizing activity.
10. Which statement about the homeobox is *not* true?
 - a. It is transcribed and translated.
 - b. It is found only in animals.
 - c. Proteins containing the homeodomain bind to DNA.
 - d. It is a sequence of DNA shared by more than one gene.
 - e. It occurs in Hox genes.

FOR DISCUSSION

1. Molecular biologists can attach genes to active promoters and insert them into cells (see Section 18.5). What would happen if the following were inserted and overexpressed? Explain your answers.
 - a. *ced-9* in embryonic neuron precursors of *C. elegans*
 - b. *MyoD* in undifferentiated myoblasts
 - c. the gene for Sonic hedgehog in a chick limb bud
 - d. *Nanos* at the anterior end of the *Drosophila* embryo
2. A powerful method to test for the function of a gene in development is to generate a “knockout” organism, in which the gene in question is inactivated (see Section 18.4). What do you think would happen in each of the following cases?
 - a. a knocked-out *ced-9* in *C. elegans*
 - b. a knocked-out *Nanos* in *Drosophila*
3. If you wanted a rose flower with only petals, what kind of homeotic mutation would you seek in the rose genome?
4. During development, an animal cell’s potential for differentiation becomes ever more limited. In the normal course of events, most cells in the adult animal have the potential to be only one or a few cell types. On the basis of what you have learned in this chapter, discuss possible mechanisms for the progressive limitation of the cell’s potential.
5. How were biologists able to obtain such a complete accounting of all the cells in *C. elegans*? What major conclusions came from these studies?

ADDITIONAL INVESTIGATION

Cloning involves considerable reprogramming of gene expression in a differentiated cell so that it acts like an egg cell.

How would you investigate this reprogramming?

WORKING WITH DATA (GO TO yourBioPortal.com)

Cloning a Mammal In this hands-on exercise, you will examine the experimental protocol used by Wilmut and colleagues to clone Dolly the sheep (Figure 19.4). You will see the data on

the efficiency of this process, as well as the genetic evidence that Dolly was indeed a clone.

The eyes have it

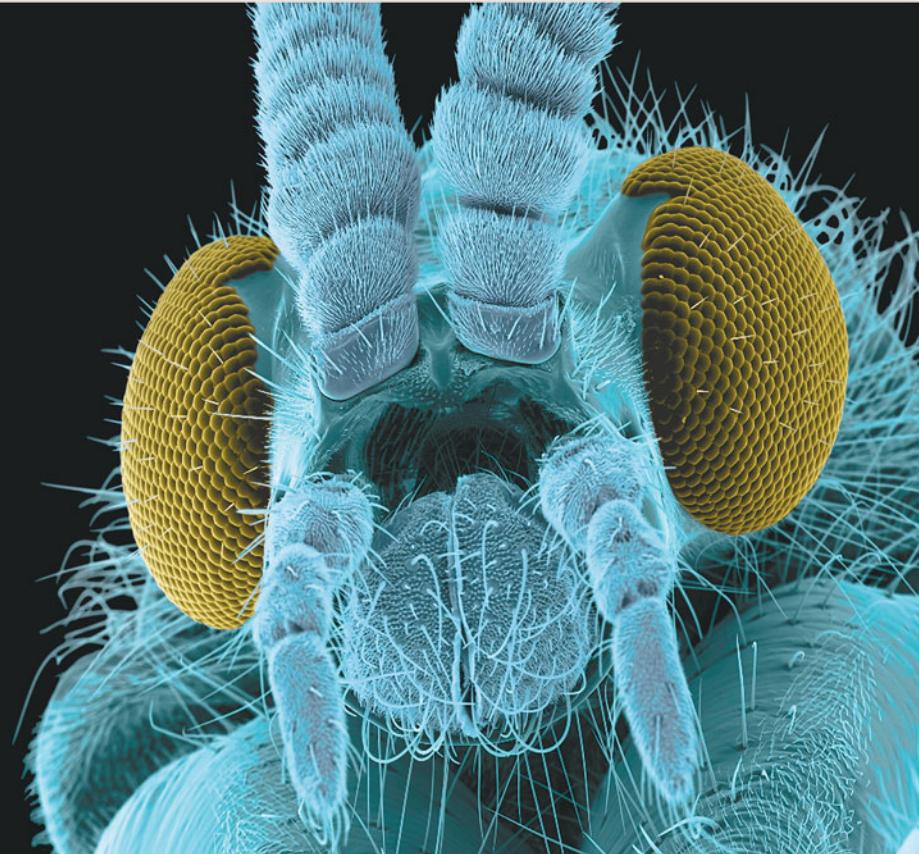
Eyes are not essential for survival; many animals and all plants get by just fine without them. However, almost all animals *do* have eyes or some type of light-sensing organs, and having eyes can confer a selective advantage.

About a dozen different kinds of eyes are found among the different animals, including the camera-like eyes of humans and the compound eyes of insects, with their thousands of individual units. In trying to understand the origin of this variety, scientists—starting with Charles Darwin—proposed that eyes evolved independently many times in different animal groups, and that each improvement in the ability of eyes to gather light and form images conferred a selective advantage on their possessor.

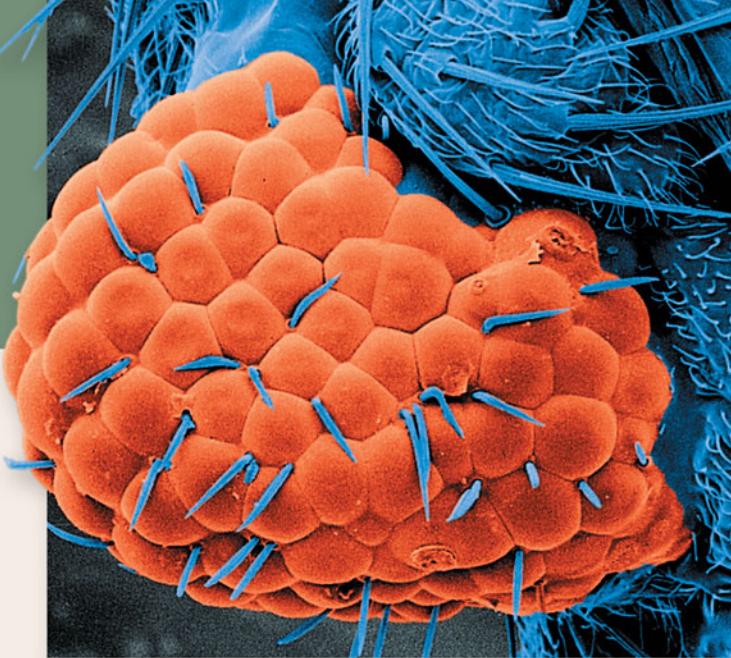
A remarkable discovery in the 1990s may have overturned this long-held dogma about the evolution of eyes. Years earlier a mutant fruit fly without eyes was found,

and the gene involved—appropriately called *eyeless*—was mapped onto one of its chromosomes. This mutant fly remained a laboratory curiosity until 1994, when the Swiss developmental biologists Rebecca Quiring and Walter Gehring began looking for transcription factors that might be involved in fly development. The gene for one of the proteins they identified mapped to the *eyeless* locus. Thus, the product of the *eyeless* gene is a transcription factor that controls the formation of the eye. Quiring and Gehring demonstrated this by making recombinant DNA constructs that allowed the *eyeless* gene to be expressed in various embryonic tissues of transgenic flies. These experiments resulted in adult flies with extra eyes on various body parts—including on the legs, under the wings, and on the antennae—depending on where the *eyeless* gene was expressed in the embryos.

But the big surprise came when the scientists performed a database search and found that the *eyeless* gene sequence was quite similar to that of *Pax6*, a gene in mice that, when mutated, leads to the development of very small eyes. Could the very different eyes of flies and mice just be variations on a common developmental theme? To test for functional similarity between the insect and mammalian genes, Gehring and colleagues repeated their gene expression experiments using the mouse *Pax6* gene instead of the fly *eyeless* gene. Once again, eyes developed at various sites on the transgenic flies. So a gene whose expression normally leads to the development of a mammalian “camera” eye now led to the development of an insect’s “compound”



Eye of the Fly Unlike the single-lensed eyes of vertebrates, the compound eyes of flies and other insects are composed of thousands of individual lenses, or ommatidia.



A Mouse Gene Can Produce a Fly's Eye When the mouse Pax6 eye-specifying gene was implanted in the part of the fruit fly embryo that normally produces a limb, ommatidia emerged in place of a leg.

eye—a very different eye type. Thus a single transcription factor appears to function as a molecular switch that turns on eye development. Although eyes evolved many times during animal evolution, all of them depend on the same gene. The special features of the many different eyes in diverse animals all evolved from a common developmental process.

The discovery that the same genes govern development in a wide variety of animals led to the rapid growth of the discipline of evolutionary developmental biology, often known as “evo-devo.” Evolutionary developmental biologists compare the genes that regulate development in many different multicellular organisms to understand how a single gene can do so many different things.

IN THIS CHAPTER we show that the genes controlling pattern formation, which we introduced in Chapter 19, are shared by a diverse array of organisms. We next describe how changes can occur in some parts of an organism without causing undesirable changes in other parts. We see how a common set of genes can produce a great variety of body forms. We then turn to the ways some organisms can modulate their development by responding to signals from their environment. Finally, we examine how developmental processes constrain evolution.

CHAPTER OUTLINE

- 20.1** What Is Evo-Devo?
- 20.2** How Can Mutations with Large Effects Change Only One Part of the Body?
- 20.3** How Can Differences among Species Evolve?
- 20.4** How Does the Environment Modulate Development?
- 20.5** How Do Developmental Genes Constrain Evolution?

20.1 What Is Evo-Devo?

The modern study of evolution and development is called **evo-devo**, or **evo-devo**. Its ideas have come from studies of the molecular mechanisms that underlie the development of morphology, and how the genes controlling these mechanisms have evolved. The principles of evo-devo are:

- Many groups of animals and plants, even distantly related ones, share similar molecular mechanisms for morphogenesis and pattern formation. As we saw in the opening essay, some genes that are experimentally swapped from one organism to another can retain similar functions in the new organisms. These mechanisms can be thought of as “toolkits,” in the same sense that a few tools in a carpenter’s toolkit can be used to build many different structures.
- The molecular pathways that determine different developmental processes, such as anterior–posterior polarity and organ formation in animals, operate independently from one another. This is called **modularity**.
- Changes in the location and timing of expression of particular genes are important in the evolution of new body forms and structures.
- Development produces morphology, and much of morphological evolution occurs by modifications of existing development genes and pathways, rather than the introduction of radically new developmental mechanisms.
- Mechanisms of development have often evolved to be responsive to environmental conditions.

Biologists have long known that the morphological differences between species are due to differences in their genomes. But we have also discussed how the genomes of different species—including distantly related ones—share numerous similar regulatory and coding sequences (see Section 17.3). When developmental biologists began to describe the events of differentiation, morphogenesis, and pattern formation at the molecular level, they found common regulatory genes and pathways in organisms that don’t appear similar at all, such as fruit flies and mice.

Developmental genes in distantly related organisms are similar

In the opening story of this chapter, we describe how a single developmental switch turns on the production of eyes in two widely divergent species—fruit flies and mice—that are only



Mouse *Pax6* gene:

DNA	GTATCCAACGGTTGTGAGTAAAATTCTGGGCAGGTATTACGAGACTGGCTCCATCAGA
Amino acids	V S N G C V S K I L G R Y Y E T G S I R



Fly *eyeless* gene:

77%	GTATCAAATGGATGTGTGAGCAAAATTCTCGGGAGGTATTATGAAACAGGAAGCATACGA
100%	V S N G C V S K I L G R Y Y E T G S I R



Shark eye control gene:

85%	GTGTCCAACGGTTCTGTCAAGTAAAATCCTGGGCAGATACTATGAAACAGGAATCCATCAGA
100%	V S N G C V S K I L G R Y Y E T G S I R



Squid eye control gene:

78%	GTCTCCAACGGCTGCGTTAGCAAGATTCTCGGACGGTACTATGAGACGGGCTCCATAAGA
100%	V S N G C V S K I L G R Y Y E T G S I R

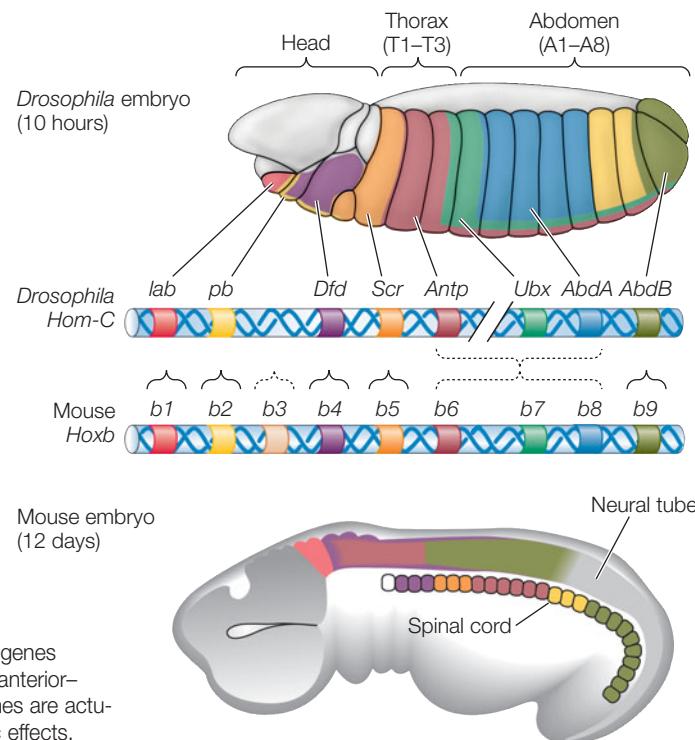
20.1 DNA Sequence Similarity in Eye Development Genes Genes controlling eye development contain regions that are highly conserved, even among species with very different eyes. These sequences, from a conserved region of the *Pax6* gene and its homologs in other species, are similar at the DNA level (top sequence in each pair) and identical at the amino acid level (bottom sequence). The percentages beside the sequences represent the percent match with the corresponding DNA and protein sequences in the mouse.

distantly related by evolution. The genes that control this switch, *eyeless* in fruit flies and *Pax6* in mice, contain sequences that are highly conserved in these species and in other animals (Figure 20.1). As described in Section 22.1, biologists infer from these similarities that the genes are *homologous*, meaning that they evolved from a gene present in a common ancestor of mice and fruit flies. In recent years, thousands of genes have been found that are homologous across distantly related species.

An even more dramatic example of homology in genes that control development, because it involves a whole set of genes, is the Hox gene cluster. These genes provide positional information and control pattern formation in early *Drosophila* embryos (see Figure 19.19). When scientists looked for similar sequences in the mouse and human genomes, the results were amazing. The Hox genes had homologs in mammals, and what is more, the genes were arranged in similar clusters in the genomes of mammals and fruit flies, and were expressed in similar patterns in their embryos (Figure 20.2). Over the millions of years that have elapsed since the common ancestor of these animals, the genes in question have mostly been main-

tained, suggesting that their functions were favored over many different conditions.

These and other examples have led evo-devo biologists to the idea that certain developmental mechanisms, controlled by specific DNA sequences, have been conserved over long periods during the evolution of multicellular organisms. These sequences comprise the **genetic toolkit**, which has been modified and reshuffled over the course of evolution to produce the remarkable diversity of plants, animals, and other organisms in the world today.



20.2 Regulatory Genes Show Similar Expression Patterns Homologous genes encoding similar transcription factors are expressed in similar patterns along the anterior-posterior axes of both insects and vertebrates. The mouse (and human) Hox genes are actually present in multiple copies; this prevents a single mutation from having drastic effects.

20.1 RECAP

Changes in development underlie the changes in morphology that result in evolution of body form. Evo-devo is the study of the evolutionary aspects of development. Evo-devo uses the knowledge gained from, and techniques of, molecular genetics. A genetic toolkit consisting of highly conserved regulatory genes encoding transcription factors governs pattern formation in multicellular organisms.

- Describe the main ideas of evo-devo. See p. 427
- How does the story of the eye-determining genes *Pax6* and *eyeless* demonstrate the existence of common developmental genes in different organisms? See pp. 427–428 and Figure 20.1
- What is the significance of the similarity between the fruit fly and mouse genes involved in the development of polarity and pattern formation? See p. 428 and Figure 20.2

Many developmental mutations in fruit flies that result in striking abnormalities (e.g., a head segment that forms a leg; see Figure 19.20) affect only a single structure, segment or region. The rest of the embryo is often unaffected. How is this possible?

20.2 How Can Mutations With Large Effects Change Only One Part of the Body?

In Chapter 19 we describe how development involves interactions between gene products, which determine a sequence of transcriptional events leading to differential gene expression. On the other hand, the study of homeotic mutations revealed that embryos, like adults, are made up of **developmental modules**—functional entities encompassing genes and various signaling pathways that determine physical structures such as body segments and legs.

The form of each module in an organism may be changed independently of the other modules because some developmental genes exert their effects on only a single module. For example, the form of a developing animal's heart can change independently of changes in its limbs, because some of the genes that govern heart formation do not affect limb formation, and vice versa. If this were not true, a mutation in any developmental gene might result in an adult with multiple, widely different deformities. Such an adult would have difficulty functioning well in any environment.

yourBioPortal.com
GO TO Animated Tutorial 20.1 • Modularity

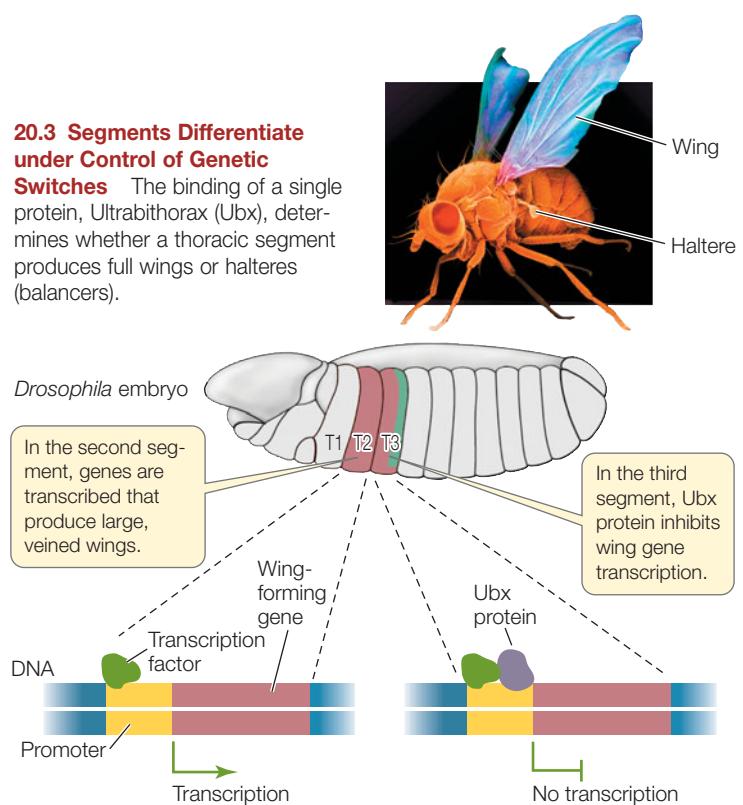
Genetic switches govern how the genetic toolkit is used

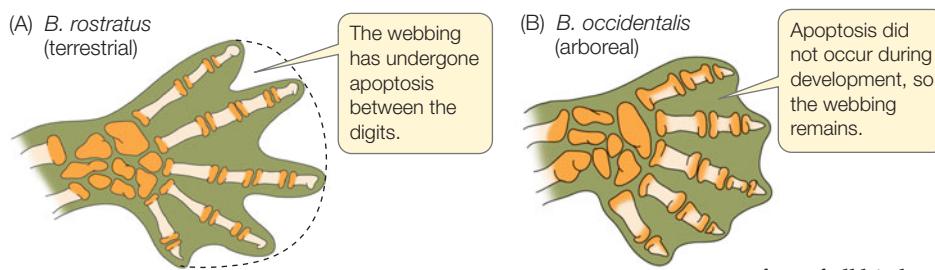
Different structures can evolve within a single organism using a common set of genetic instructions because there are mechanisms called **genetic switches** that control how the genetic toolkit is used.

These mechanisms involve promoters and the transcription factors that bind them. The signal cascades that converge on and operate these switches determine when and where genes will be turned on and off. Multiple switches control each gene by influencing its expression at different times and in different places. In this way, elements of the genetic toolkit can be involved in multiple developmental processes while still allowing individual modules to develop independently.

Genetic switches integrate positional information in the developing embryo and play key roles in determining the developmental pathways of different modules. For example, each Hox gene codes for a transcription factor that is expressed in a particular segment or appendage of the developing fruit fly. The pattern and functioning of each segment depend on the unique Hox gene or combination of Hox genes that are expressed in the segment.

Consider the formation of fruit fly wings. *Drosophila* has three thoracic segments, the first of which bears no wings. The second segment bears the large forewings, and the third segment bears small hindwings, called *halteres*, that function as balancing organs. Hox proteins are not expressed in forewing cells, but all hind wing cells express the Hox gene *Ultrabithorax* (*Ubx*) because a set of genetic switches activates the *Ubx* gene in the third thoracic segment. *Ubx* turns off genes that promote the formation of the veins and other structures of the forewing, and it turns on genes that promote the formation of hind wing features (Figure 20.3). In butterflies, on the other hand, *Ubx* influences target genes so that wings develop in the third-segment cells, so full hind wings develop. Therefore, a simple genetic change results in a major morphological difference in the wings of flies and butterflies.





Modularity allows for differences in the timing and spatial pattern of gene expression

Modularity allows the relative *timing* of different developmental processes to shift independently of one another, in a process called **heterochrony**. That is, the genes regulating the development of one module (say, the eyes of vertebrates) may be expressed at different developmental stages in different species.

Salamanders of the genus *Bolitoglossa* illustrate how heterochrony can result in major morphological changes. Salamander embryos have webbing between their toes, but in most species of salamanders a particular gene triggers apoptosis in the webbing as the salamanders develop. The resulting independent digits allow the adult salamander to walk more easily than if it had webbed feet. This is the case with *Bolitoglossa rostratus*, a species that lives on the forest floor (Figure 20.4A). But in arboreal species such as *Bolitoglossa occidentalis* (Figure 20.4B), this gene is not expressed and apoptosis does not occur. The feet of *B. occidentalis* are webbed throughout life, acting like suction cups so the animal can adhere to vertical surfaces such as tree trunks. Thus a simple change in gene expression led to a major morphological change and allowed a new lifestyle.

The evolution of the giraffe's neck provides another example of heterochrony. As in virtually all mammals (with the exception of manatees and sloths) there are seven vertebrae in the neck of the giraffe. So the giraffe did not get a longer neck by adding vertebrae. Instead the cervical (neck) vertebrae of the giraffe are much longer than those of other mammals (Figure 20.5). Bones grow due to the proliferation of cartilage-producing cells called chondrocytes. Bone growth is stopped by a signal that results in death of the chondrocytes and calcification of the bone matrix. In giraffes this signaling process is delayed in the cervical vertebrae, with the result that these vertebrae grow longer. Thus, the evolution of longer necks acted through *changes in the timing of expression* of the genes that control bone formation.

Differences in the *spatial expression pattern* of a developmental gene can also result in evolutionary change. Foot webbing in salamanders is determined by the temporal expression of a developmental gene, but foot webbing in ducks and chickens is affected by alterations in the spatial expression of a gene. The

20.4 Heterochrony Resulted in the Evolution of a Tree-Climbing Salamander

(A) The foot of an adult *B. rostratus*, a terrestrial salamander. (B) The foot of *B. occidentalis*, a closely related salamander, does not lose its webbing. This species uses the suction of its webbed feet in an arboreal lifestyle.

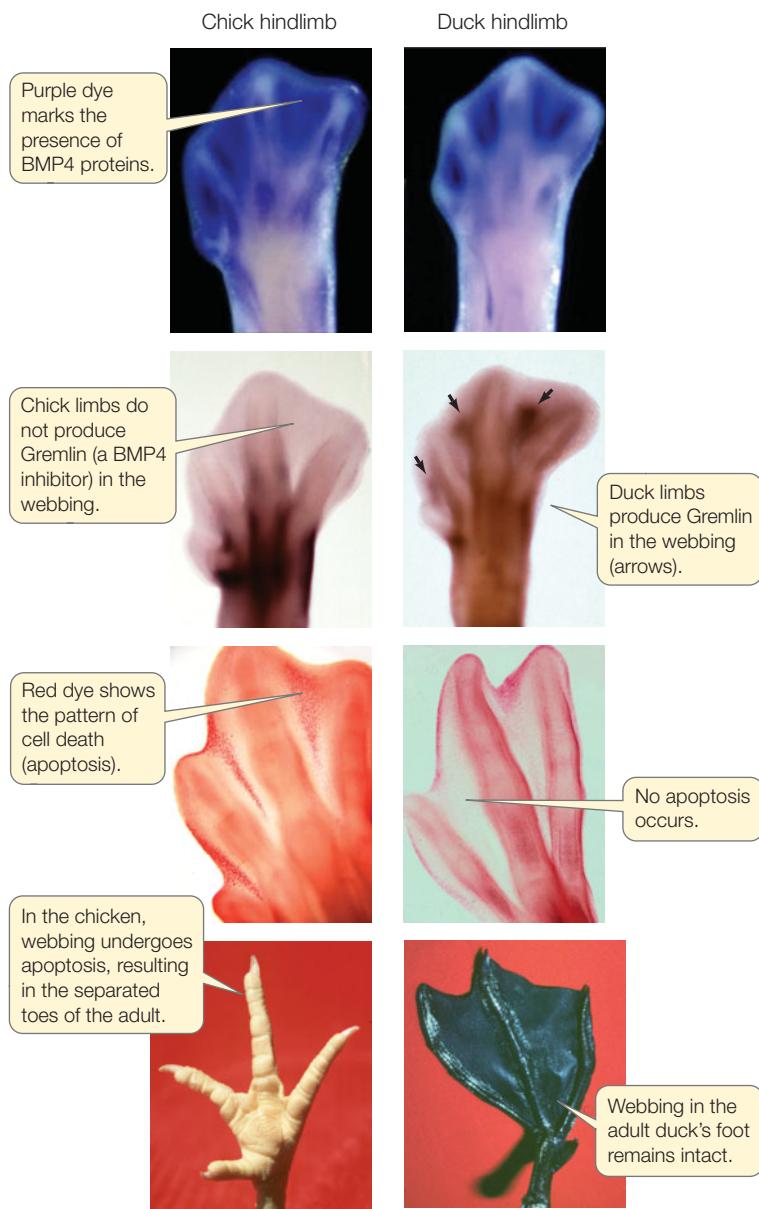
feet of all bird embryos have webs of skin that connect their toes. This webbing is retained in adult ducks (and other aquatic birds) but not in adult chickens (and other non-aquatic birds). The loss of webbing is due to a signaling protein called bone morphogenetic protein 4 (BMP4) that instructs the cells in the webbing to undergo apoptosis. The death of these cells destroys the webbing between the toes.

Embryonic duck and chicken hindlimbs both express the BMP4 gene in the webbing between the toes, but they differ in expression of a gene called *Gremlin*, which encodes a BMP *inhibitor* protein (Figure 20.6). In ducks, but not chickens, the *Gremlin* gene is expressed in the webbing cells. The *Gremlin* protein inhibits the BMP4 protein from signaling for apoptosis, and the result is a webbed foot. If chicken hindlimbs are experimentally exposed to *Gremlin* during development, apoptosis does not occur, and ducklike webbed feet form on the chicken (Figure 20.7).

(A) Giraffe



20.5 Heterochrony in the Development of a Longer Neck There are seven vertebrae in the neck of the giraffe (left) and human (right; not to scale). But the vertebrae of the giraffe are much longer (25 cm compared to 1.5 cm) because during development, growth continues for a longer period of time. This timing difference is called heterochrony.

**20.6 Changes in Gremlin Expression Correlate with Changes in**

Hindlimb Structure The left column of photos shows the development of a chicken's foot; the right column shows foot development in a duck. Gremlin protein in the webbing of the duck foot inhibits BMP4 signaling, thus preventing the embryonic webbing from undergoing apoptosis.

20.2 RECAP

Embryos and adult organisms are made up of self-contained units called modules. The form of each module may change independently because some developmental genes exert their effects on only a single module.

- How do genetic switches control the way a gene is used? See p. 429 and Figure 20.3
- Explain how heterochrony can result in evolutionary change. See p. 430 and Figures 20.4 and 20.5

INVESTIGATING LIFE**20.7 Changing the Form of an Appendage**

Ducks have webbed feet and chickens do not—a major difference in the adaptations of these species. Webbing is initially present in the chick embryo, but undergoes apoptosis that is stimulated by the protein BMP4. In ducks another protein, Gremlin, binds to BMP4 and inhibits it, preventing apoptosis and resulting in webbed feet. J. J. Hurle and colleagues at the Universidad de Cantabria in Spain asked what would happen if Gremlin were put onto a developing chick foot. They hypothesized that apoptosis would be inhibited, and it was: the chick developed webbed feet. Thus, a single developmental switch controls foot shape—an important adaptation to the environment.

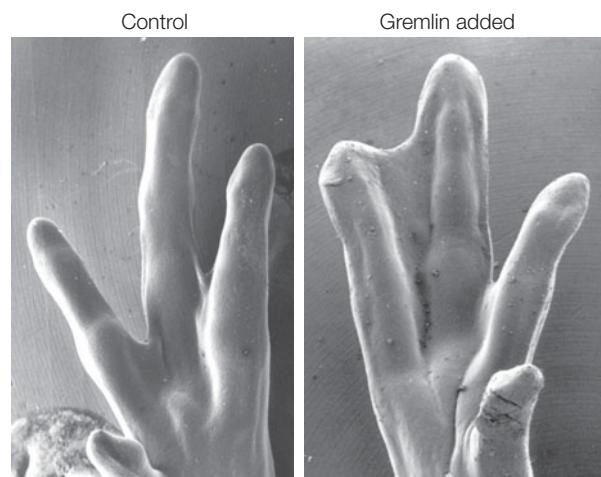
HYPOTHESIS Adding Gremlin protein (a BMP4 inhibitor) to a developing chicken foot will transform it into a ducklike foot.

METHOD

Chip a small window in chick egg shell and carefully add Gremlin-secreting beads to the webbing of embryonic chicken hindlimbs. Add beads that do not contain Gremlin to other hindlimbs (controls). Close the eggs and observe limb development.

RESULTS

In the hindlimbs in which Gremlin was secreted, the webbing does not undergo apoptosis, and the hindlimb resembles that of a duck. The control hindlimbs develop the normal chicken form.

**CONCLUSION**

Differences in *Gremlin* gene expression cause differences in morphology, allowing duck hindlimbs to retain their webbing.

Go to yourBioPortal.com for original citations, discussions, and relevant links for all INVESTIGATING LIFE figures.

Genetic manipulations and studies of pattern formation within embryos have shown that the same signals can control development of different structures in an individual organism. For example, the protein BMP4 promotes apoptosis between developing digits in feet, and then is involved later in the formation

of bone. These studies suggest that the processes that generate multiple structures *within* an organism might also explain how different structures develop in different species.

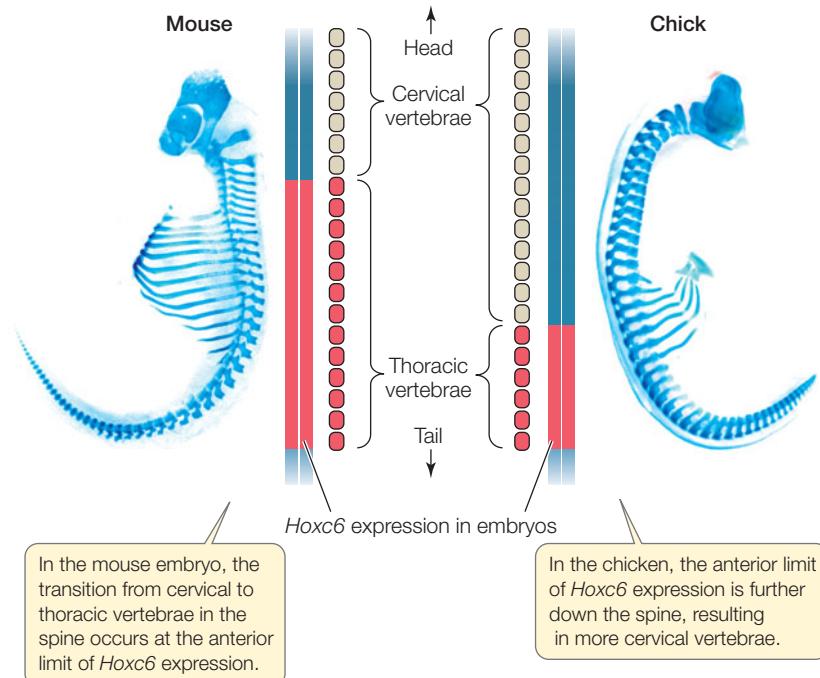
20.3 How Can Differences among Species Evolve?

Can the processes that allow different structures to develop in different regions of an embryo also explain major morphological differences among species? Apparently they can. The genetic switches that determine where and when genes will be expressed appear to underlie both the transformation of an individual from egg to adult and the many major differences in body form that exist among species. Arthropods provide good examples of how morphological differences among species can evolve through mutations in the genes that regulate the differentiation of segments.

The arthropods (which include crustaceans, centipedes, spiders, and insects) are segmented, with head, thoracic, and abdominal segments. In centipedes, both thoracic and abdominal segments form legs; but in insects, only the thoracic segments do. Arthropods express a gene called *Distal-less* (*Dll*) that causes legs to form from segments. What shuts down *Dll* expression in insect abdominal segments? The product of the Hox gene *Ubx* is produced in arthropod abdominal segments. But it has very different effects in different organisms. In centipedes, the *Ubx* protein apparently activates expression of the *Dll* gene to promote the formation of legs. But during the evolution of insects, a change in the *Ubx* gene sequence resulted in a modified *Ubx* protein that *represses* *Dll* expression in abdominal segments, so leg formation is inhibited. A phylogenetic tree of arthropods shows that this change in *Ubx* occurred in the ancestor of insects, at the same time that abdominal legs were lost (Figure 20.8).

20.8 A Mutation in a Hox Gene

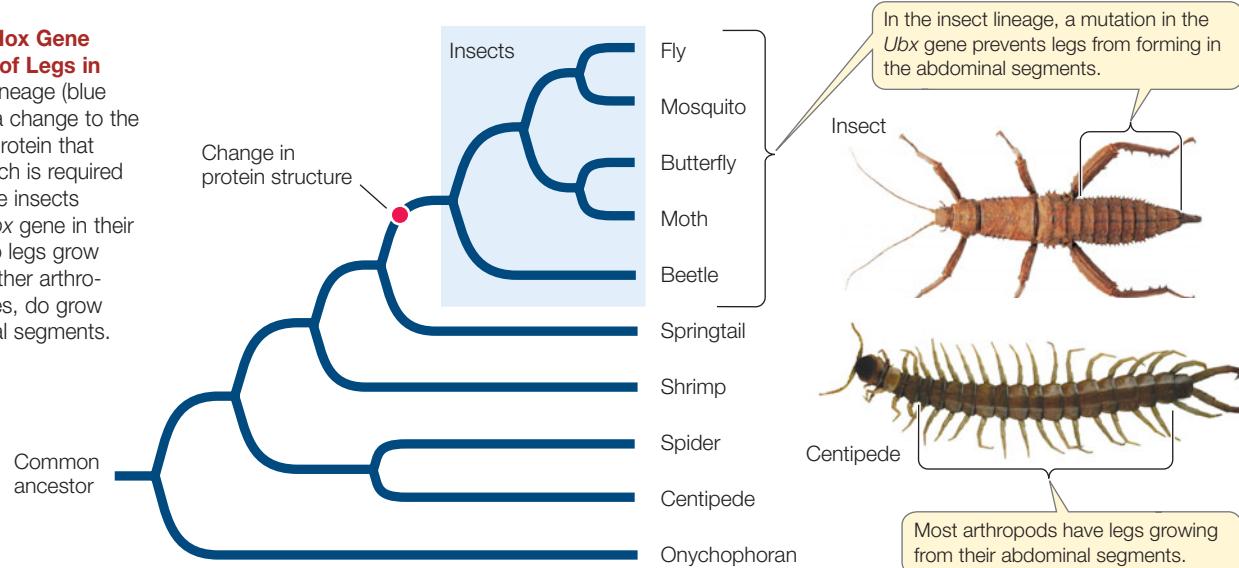
Changed the Number of Legs in Insects In the insect lineage (blue box) of the arthropods, a change to the *Ubx* gene resulted in a protein that inhibits the *Dll* gene, which is required for legs to form. Because insects express this modified *Ubx* gene in their abdominal segments, no legs grow from these segments. Other arthropods, such as centipedes, do grow legs from their abdominal segments.



20.9 Changes in Gene Expression and Evolution of the Spine

Differences in the pattern of *Hoxc6* expression result in a different boundary between the cervical and thoracic vertebrae in mice and chicks.

In vertebrates, a similar process governs the development of differences in segments of the vertebral column. The vertebral column consists of a set of anterior-to-posterior regions: the cervical (neck), thoracic (chest), lumbar (back), sacral, and caudal (tail) regions. The spatial pattern of Hox gene expression governs the transition from one region to another (Figure 20.9). For example, the anterior limit of expression of *Hoxc6* always falls at the boundary between the cervical and thoracic vertebrae of mice and chickens, even though these animals have different numbers of cervical and thoracic vertebrae. The anterior-most



segment that expresses *Hoxc6* is the segment where the forelimbs will develop. Thus, genetic changes that expanded or contracted the expression domains of different Hox genes resulted in changes in the characteristic numbers of different vertebrae during evolution.

20.3 RECAP

Changes in the genetic switches that determine where and when genes will be expressed underlie the evolution of differences in form among species.

- Why do insects, unlike other arthropods, lack abdominal limbs? See p. 432 and Figure 20.8
- How can the evolution of the spinal column be explained by changes in a developmental gene? See p. 432 and Figure 20.9

So far in this chapter, we have focused on how modular genetic signaling cascades control the development of an organism and how changes in genetic switches can produce differences between species. You may have the impression that all of these processes unfold from the genetic information contained in the fertilized egg, but that is not the case. Information from the environment can influence the genetic signaling cascades and thereby alter the form of the organism.

20.4 How Does the Environment Modulate Development?

The environment an individual lives in may differ from the one its parents lived in. Some environmental signals can produce developmental changes in an organism. If such changes result in higher reproductive fitness, they will be favored by natural selection. The ability of an organism to modify its development in response to environmental conditions is called **developmental plasticity** or *phenotypic plasticity*. It means that *a single genotype has the capacity to produce two or more different phenotypes*.

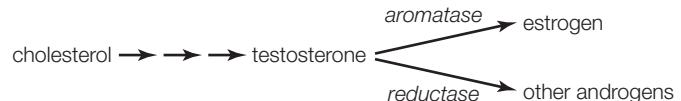
Temperature can determine sex

In Chapter 12 we discuss the genetic mechanisms that determine sex. In mammals there are two sex chromosomes; XX individuals are female and XY individuals are male. But in some reptiles, sex is determined not by genetic differences between individuals, but rather by the temperature at which the eggs are incubated—a remarkable case of developmental plasticity.

Research in the laboratory of David Crews at the University of Texas has shown that if eggs of the red-eared slider turtle are incubated at temperatures below 28.6°C, they will all become males, whereas if the eggs are incubated above 29.4°C, they will all become females. In the less than 1°C range between these two temperatures, a mix of males and females will hatch from the eggs (Figure 20.10). In other species with temperature-dependent sex determination, the incubation temperatures that produce males and females may differ from those that produce

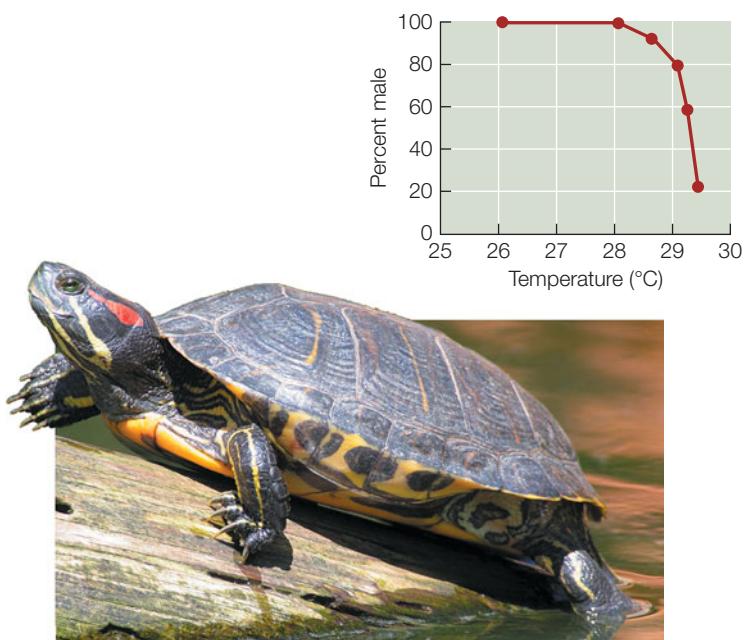
males and females in the red-eared slider. These different temperature dependencies indicate that the effects of incubation temperature can vary among species. But how does temperature control this developmental plasticity?

In vertebrates, the development of male and female organs in the embryo is controlled by the actions of sex steroid hormones. This is the case whether the organism's sex determination is controlled genetically or by temperature. Sex steroid biosynthesis in both males and females begins with cholesterol, which goes through many chemical reactions to produce the male sex steroids (androgens) and the female sex steroids (estrogens). In this biosynthetic sequence, the step that produces the first androgen—testosterone—precedes the step that produces estrogens; therefore, both males and females produce testosterone.



In animals with temperature-controlled sex determination, incubation temperature influences sex development by controlling the expression of the enzyme aromatase, which converts testosterone to estrogen. If aromatase is abundantly expressed, estrogens are dominant and female organs develop. If aromatase is not expressed, testosterone is dominant, and male organs develop. Applying estrogen to eggs results in the development of females, even at the male-inducing temperature.

What is the evolutionary advantage of this sex determination mechanism? It might be that incubation temperature influences the growth rate of the embryo and the time of hatching.



20.10 Hot Females, Cool Males Whether the embryo of a red-eared slider turtle develops into a male or a female depends on the temperature at which the egg is incubated. Higher temperatures produce only females and lower temperatures produce only males. This is apparently due to temperature sensitivity in the synthesis of sex hormones.

In species in which males compete for territories and for females, a larger body size would be a benefit for males, but not necessarily for females. Depending on availability of food in the environment, earlier hatching may have a positive or a negative effect on growth rate. For these kinds of reasons, incubation temperature may have a differential effect on the reproductive successes of males and females in a population.

One interesting experiment clearly demonstrated the fitness value of temperature-determined sex (Figure 20.11). At the University of Sydney, Daniel Warner and Richard Shine used hormones to manipulate the eggs of a lizard called the jacky dragon (*Amphibolurus muricatus*) to produce males at temperatures that would normally result in females. In this species, females are produced at all incubation temperatures, but males are only produced at incubation temperatures between 27°C and 30°C. The hormonal manipulations allowed the investigators to obtain both males and females from three different incubation temperatures—low, medium, and high—and to compare their subsequent growth characteristics and reproductive success.

The young lizards were released into outdoor enclosures and allowed to behave naturally for the next three years. The males incubated at the medium temperature had higher reproductive success over the three-year study. The reproductive successes of females from the low and medium incubation temperatures did not differ, but were higher for the high-temperature female group. These data support the hypothesis that the incubation temperature differentially affects reproductive success in males and females, and provides an explanation for why there could be selection for temperature-dependent sex determination in some species.

Organisms use information that predicts future conditions

In many cases of phenotypic plasticity, the adaptation of the different body forms to different but predictable environments is quite obvious. An excellent example is the moth *Nemoria arizonaria*, which produces two generations each year. Caterpillars

INVESTIGATING LIFE

20.11 Temperature-Dependent Sex Determination is Associated with Sex-Specific Fitness Differences

In some reptiles, sex is determined by the incubation temperature of the developing embryo. This led to the hypothesis that male-inducing temperatures during development result in males with higher reproductive fitness. Warner and Shine tested this hypothesis by using a drug to block estrogen synthesis, so that males developed instead of females at high and low temperatures. These males had much lower reproductive fitness than males that developed at the normal male-inducing temperature (which is intermediate). In contrast, females showed highest fitness when they developed from eggs incubated at higher temperatures.

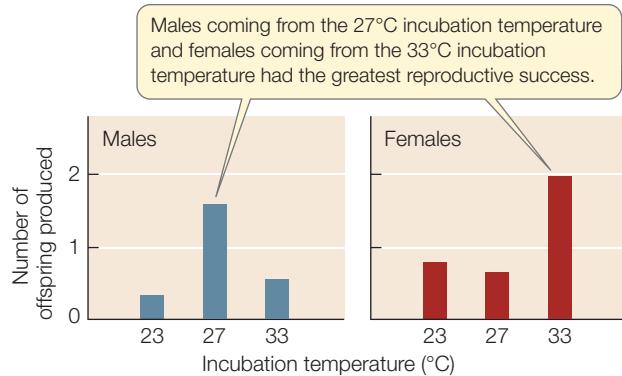
HYPOTHESIS Incubation temperature has a differential effect on reproductive success in lizards.

METHOD

Incubate jacky dragon eggs at 23°C, 27°C, and 33°C. Apply aromatase inhibitor to half the eggs. Raise lizards in the natural environment and record the number of offspring produced.

RESULTS

Untreated eggs: males at 27°C; females at 23°C and 33°C
Inhibitor-treated eggs: males at 23°C, 27°C, and 33°C



CONCLUSION

Incubation temperature during development has a differential effect on fitness of male and female jacky dragons, and thus could be a selective pressure leading to the pattern of temperature determination of sex in this species.

that hatch from eggs in spring feed on oak tree flowers (catkins). These caterpillars complete their development and transform into adult moths in summer. The summer moths lay their eggs on oak leaves, and the caterpillars that hatch eat the leaves. When these caterpillars transform into adult moths, they lay eggs that overwinter and hatch in the spring when the catkins are once again in bloom. Both types of caterpillars are camouflaged in the environments in which they feed. The body form of the spring caterpillars resembles the catkins (**Figure 20.12A**), and the body form of the summer caterpillars resembles small oak branches (**Figure 20.12C**). At the time of hatching, the young caterpillars all look similar, but their diets trigger developmental changes that result in the differences in appearance. The ability to avoid predation by phenotypic plasticity increases evolutionary fitness.

A variety of environmental signals influence development

In addition to temperature and diet, there are other environmental signals that initiate developmental changes. A ubiquitous and dependable source of environmental information is light, which provides predictive information about seasonal changes. Outside of the equatorial region, lengthening days herald spring and summer while shortening days indicate oncoming fall and winter. Many insects use day length to enter or exit a period of developmental or reproductive arrest called **diapause**, which enables them to better survive harsh conditions. Deer, moose, and elk use day length to time the development and the dropping of antlers, and many organisms use day length to optimize the timing of reproduction. As we discuss in Chapter 38, many plants initiate reproduction in response to the length of the night (an absence of light) and others respond to certain wavelengths of light with developmental changes.

You may wonder why we are mentioning processes like antler growth and seasonal reproduction in a chapter on development. Development encompasses more than the events that occur before an organism reaches maturity.

Development includes changes in body form and function that can occur throughout the life of the organism.

Plants provide a particularly clear example of this. Redwood trees that are thousands of years old still have undifferentiated tissues called meristems that produce new differentiated tissues for the tree—stems, leaves, reproductive structures, and so on—throughout its life. These developmental processes are not a simple read-out of a genetic program; they are adjusted to optimize plant form in the environment in which the

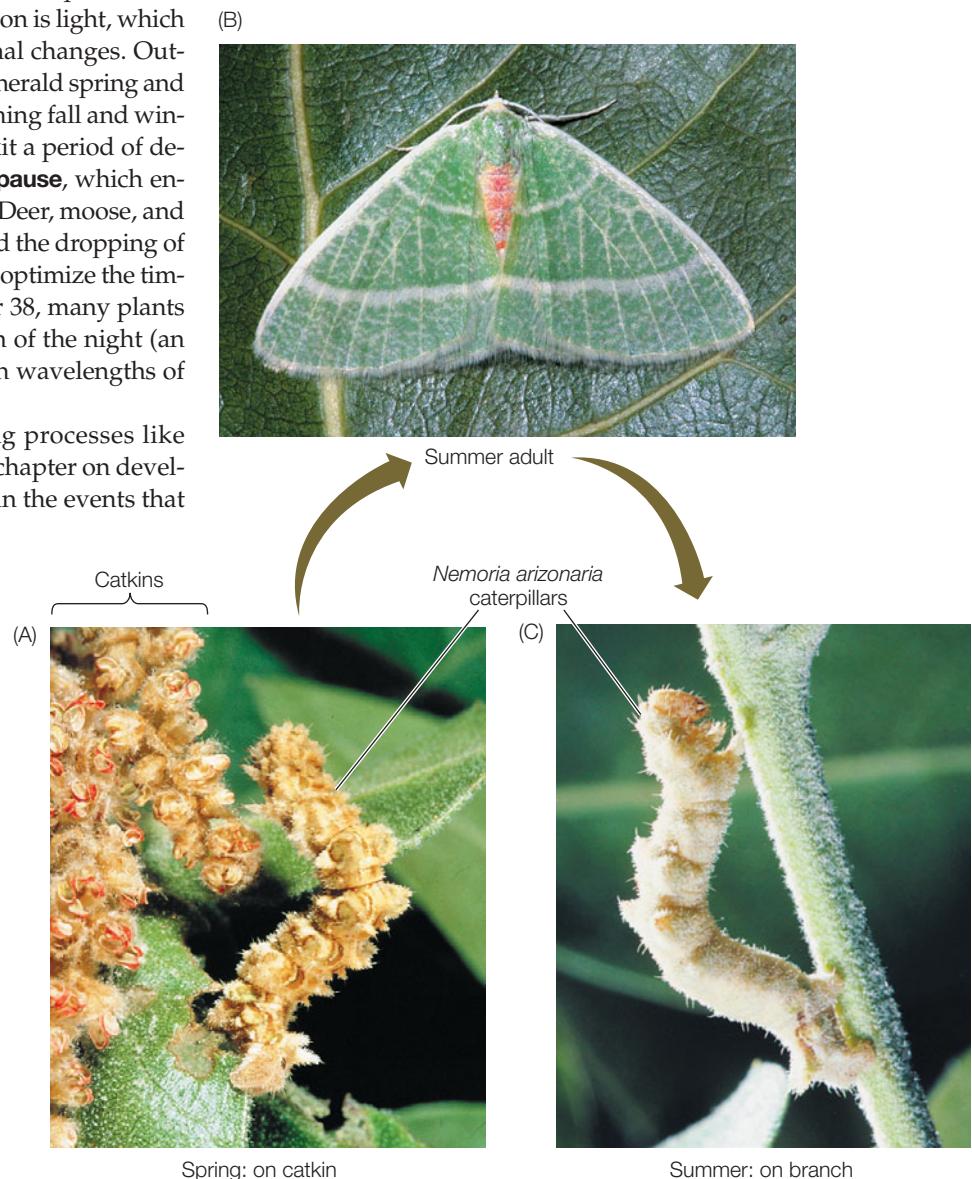
20.12 Spring and Summer Forms of a Caterpillar (A) Spring caterpillars of the moth *N. arizonaria* resemble the oak catkins on which they feed. They develop into adults (B), which lay eggs on oak leaves. The summer caterpillars (C) of the same species resemble oak twigs.

plant grows. Light, which plants need for photosynthesis, is an important environmental signal in plant development. Dim light stimulates the elongation of stem cells, so that plants growing in the shade become tall and spindly (**Figure 20.13**). This developmental plasticity is adaptive because a spindly plant is more likely to reach a patch of brighter light than a plant that remains compact. In bright light a plant does not need to grow tall, and can put its energy into growing leaves.

yourBioPortal.com

GO TO Web Activity 20.1 • Plant Development

Natural selection can act on any genes or signaling pathways with important developmental functions that can influence reproductive success. Antler growth cycles involve the turning on and off of genes controlling bone growth. Seasonal breeding involves turning on and off the same genes that were involved in sex development and maturation. The evolution of development extends to all stages of life.





20.13 Light Seekers The bean plants on the left were grown under low light levels. The plant's cells have elongated in response to the low light, and the plants have become spindly. The control plants on the right were grown under normal light conditions.

20.4 RECAP

Developmental plasticity enables developing organisms to adjust their forms to fit the environments in which they live. Organisms respond to environmental signals that are accurate predictors of future conditions. Development continues throughout life, and can result in adaptive changes in the forms and functions of adult organisms.

- Describe several examples of how an organism's phenotype can be a response to environmental signals. **See pp. 433–435 and Figures 20.10 and 20.12**
- How would you determine whether or not an environmental effect on development is adaptive? **See pp. 434–435 and Figure 20.11**

Appropriate responses to new environmental conditions are likely to evolve over time, but what are the limits of such evolution? Do developmental genes dictate what structures and forms are possible?

20.5 How Do Developmental Genes Constrain Evolution?

Four decades ago, the French geneticist François Jacob made the analogy that evolution works like a tinker, assembling new structures by *combining and modifying the available materials*, and not like an engineer, who is free to develop dramatically different designs (say, a jet engine to replace a propeller-driven engine). We have seen that the evolution of morphology has not been governed by the appearance of radically new genes, but by modifications of existing genes and their regulatory pathways. Thus, developmental genes and their expression constrain evolution in two major ways:

- Nearly all evolutionary innovations are modifications of previously existing structures.
- The genes that control development are highly conserved; that is, the regulatory genes themselves change slowly over the course of evolution.

Evolution proceeds by changing what's already there

The features of organisms almost always evolve from preexisting features in their ancestors. New “wing genes” did not suddenly appear in insects, birds, and bats; instead, wings arose as modifications of existing structures. Wings evolved independently in insects and vertebrates—once in insects, and in three independent instances among the vertebrates (Figure 20.14). In vertebrates, the wings are modified limbs.

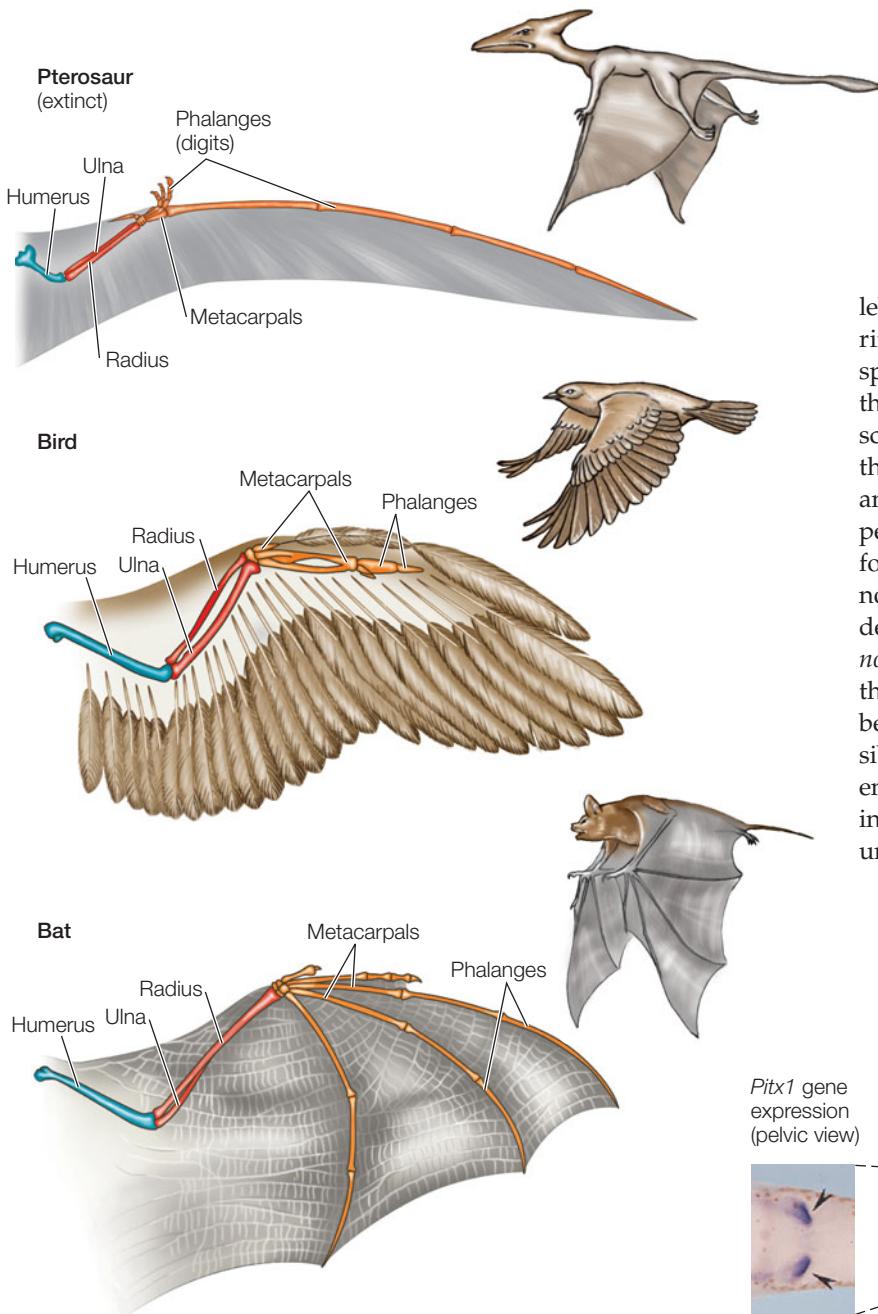
Like limbs, wings have a common structure: a humerus that articulates with the body; two longer bones, the radius and ulna, that project away from the humerus; and then metacarpals and phalanges (digits). During development these bones have different lengths and weights in different organisms.

Developmental controls also influence how organisms lose structures. The ancestors of present-day snakes lost their forelimbs as a result of changes in the segmental expression of Hox genes. The snake lineage subsequently lost its hindlimbs by the loss of expression of the *Sonic hedgehog* gene in the limb bud tissue. But some snake species such as boas and pythons still have rudimentary pelvic bones and upper leg bones.

Conserved developmental genes can lead to parallel evolution

The nucleotide sequences of many of the genes that govern development have been highly conserved throughout the evolution of multicellular organisms—in other words, these genes exist in similar form across a broad spectrum of species (see Figure 20.2).

The existence of highly conserved developmental genes makes it likely that similar traits will evolve repeatedly, especially among closely related species—a process called **parallel phenotypic evolution**. A good example of this process is provided by a small fish, the three-spined stickleback (*Gasterosteus aculeatus*).

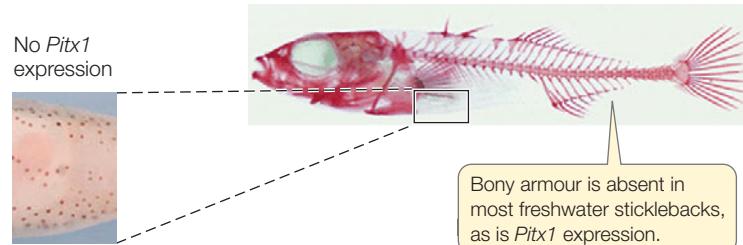
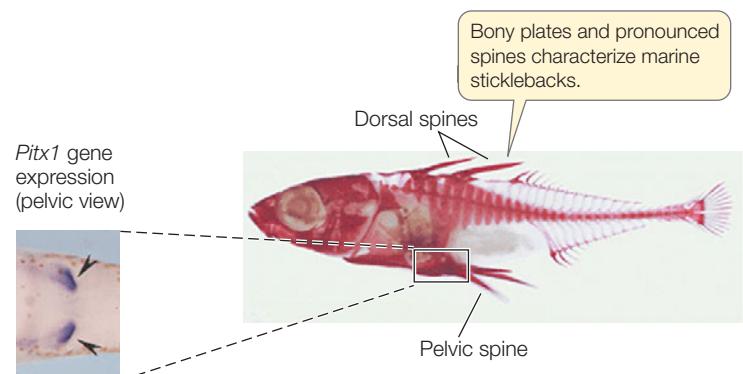


Sticklebacks are widely distributed across the Atlantic and Pacific Oceans and are also found in many freshwater lakes. Marine populations of this species spend most of their lives at sea, but return to fresh water to breed. Members of freshwater populations live in lakes and never journey to salt water. Genetic evidence shows that freshwater populations have arisen independently from marine populations many times, most recently at the end of the last ice age. Marine sticklebacks have several structures that protect them from predators: well-developed pelvic bones with pelvic spines, and bony plates. In the freshwater populations descended from them, this body armor is greatly reduced, and dorsal and pelvic spines are much shorter or even lacking (Figure 20.15).

20.14 Wings Evolved Three Times in Vertebrates

The wings of pterosaurs (the earliest flying vertebrates, which lived from 265 to 220 million years ago), birds, and bats are all modified forelimbs constructed from the same skeletal components. However, the components have different forms in the different groups of vertebrates.

The difference between marine and freshwater sticklebacks is not induced by environmental conditions. Marine species that are reared in fresh water still grow spines. Not surprisingly, the difference is due to a gene that affects development. The *Pitx1* gene codes for a transcription factor that is normally expressed in regions of the developing embryo that form the head, trunk, tail, and pelvis of the marine stickleback. However, in independent populations from Japan, British Columbia, California, and Iceland, the gene has evolved such that it is no longer expressed in the pelvis, and the spines do not develop. *This same gene has evolved to produce similar phenotypic changes in several independent populations*, and is thus a good example of parallel evolution. What could be the common selective mechanism in these cases? Possibly, the decreased predation pressure in the freshwater environment allows for increased reproductive success in animals that invest less energy in the development of unnecessary protective structures.



20.15 Parallel Phenotypic Evolution in Sticklebacks A developmental gene, *Pitx1*, encodes a transcription factor that stimulates the production of plates and spines. This gene is active in marine sticklebacks, but mutated and inactive in various freshwater populations of the fish. The fact that this mutation is found in geographically distant and isolated freshwater populations is evidence for parallel evolution.

20.5 RECAP

Developmental controls constrain evolution because nearly all evolutionary innovations are modifications of previously existing structures. The conservation of many genes makes it likely that similar traits will evolve repeatedly.

- How have diverse body forms evolved by means of modifications in the functioning of existing genes? [See p. 436 and Figures 20.14](#)
- Explain how the differences between marine and freshwater sticklebacks exemplify parallel evolution via changes in gene regulation. [See p. 437 and Figure 20.15](#)

During the course of evolution, many novel traits have arisen, but failed to persist beyond a single generation. Part Six of this book examines the processes of evolution—the powerful forces that influence the survival and reproductive success of various life forms. We will examine how different adaptations become prevalent in different environments, resulting in the extraordinary diversity of life on Earth today.

CHAPTER SUMMARY

20.1 What Is Evo-Devo?

- **Evolutionary developmental biology (evo-devo)** is the modern study of the evolutionary aspects of development, and it focuses on molecular mechanisms.
- Changes in development underlie evolutionary changes in morphology that produce major differences in body forms.
- Similarities in the basic mechanisms of development between widely divergent organisms reflect common ancestry. [Review Figure 20.1](#)
- Evolutionary diversity is produced using a modest number of regulatory genes.
- The transcription factors and chemical signals that govern pattern formation in the bodies of multicellular organisms, and the genes that encode them, can be thought of as a **genetic toolkit**.
- Regulatory genes have been highly conserved during evolution. [Review Figure 20.2](#)

20.2 How Can Mutations with Large Effects Change Only One Part of the Body?

SEE ANIMATED TUTORIAL 20.1

- The bodies of developing and mature organisms are organized into self-contained units called **developmental modules** that can be modified independently. Modularity allows the timing of different developmental processes to shift independently, in a process called **heterochrony**. [Review Figure 20.4 and 20.5](#)
- Alterations in the spatial expression patterns of regulatory genes can also result in evolutionary changes. [Review Figure 20.6](#)

20.3 How Can Differences among Species Evolve?

- Changes in **genetic switches** that determine where and when a set of genes will be expressed underlie both the transformation

of an individual from egg to adult and the evolution of differences among species.

- Morphological changes in species can evolve through mutations in the genes that regulate the differentiation of body segments. [Review Figure 20.8](#)

20.4 How Does the Environment Modulate Development?

SEE WEB ACTIVITY 20.1

- The ability of an organism to modify its development in response to environmental conditions is called **developmental plasticity**.
- In many species of reptiles, sex development is determined by incubation temperature, which acts through genes that control the production, modification, and action of sex hormones. [Review Figure 20.10](#)
- The adaptive significance of developmental plasticity is not always obvious, but experiments can test for effects on reproductive success. [Review Figure 20.11](#)
- Some environmental cues, such as those that anticipate seasons, are highly regular and can reliably drive seasonal adaptations in body form and function. [Review Figure 20.12](#)
- Environmental cues that trigger developmental change are diverse and can act at any stage of the life of an organism.

20.5 How Do Developmental Genes Constrain Evolution?

- Virtually all evolutionary innovations are modifications of pre-existing structures. [Review Figure 20.14](#)
- Because many genes that govern development have been highly conserved, similar traits are likely to evolve repeatedly, especially among closely related species. This process is called **parallel phenotypic evolution**. [Review Figure 20.15](#)

SELF-QUIZ

1. Which of the following is *not* one of the principles of evolutionary developmental biology (evo-devo)?
 - a. Animal groups share similar molecular mechanisms for morphogenesis.
 - b. Changes in the timing of gene expression are important in the evolution of new structures.
 - c. Evolution of development is not responsive to the environment.
 - d. Changes in the locations of gene expression in the embryo can lead to new structures.
 - e. Evolution occurs by modification of existing developmental genes and pathways.
2. The developmental control pathway that results in polarity and pattern formation in the head–abdomen axis in *Drosophila*
 - a. has a similar gene sequence and chromosome order in the mouse.

- b. arose only in insects during evolution.
 - c. determines only the organs that arise in head segments.
 - d. involves only gene products made by the embryo.
 - e. arose through new genes that had not existed before in any form.
3. Which of the following is *not* true of genetic switches?
- They control how a genetic toolkit is used.
 - They integrate positional information in an embryo.
 - A single switch controls each gene.
 - They allow different structures to develop within an individual organism.
 - They determine when and where a gene is turned on or off.
4. Ducks have webbed feet and chickens do not because
- ducks need webbed feet to swim, whereas terrestrial chickens do not.
 - both duck and chicken embryos express *BMP4* in the webbing between the toes, but the *Gremlin* gene is expressed in the webbing cells only in ducks.
 - both duck and chicken embryos express *BMP4* in the webbing between the toes, but the *Gremlin* gene is expressed in the webbing cells only in chickens.
 - only duck embryos express *BMP4* in the webbing between the toes.
 - only chicken embryos express *BMP4* in the webbing between the toes.
5. Modularity is important for development because it
- guarantees that all units of a developing embryo will change in a coordinated way.
 - coordinates the establishment of the anterior-posterior axis of the developing embryo.
 - allows changes in developmental genes to change one part of the body without affecting other parts.
 - guarantees that the timing of gene expression is the same in all parts of a developing embryo.
 - allows organisms to be built up one module at a time.
6. Organisms often respond to environmental signals that accurately predict future conditions by
- stopping development until the signal changes.
 - altering their development to adapt to the future environment.
- c. altering their development such that the resulting adult can produce offspring adapted to the future environment.
 - d. producing new mutants.
 - e. developing normally because the predicted conditions may not last long.
7. The process whereby changes in the timing of developmental events can change the form of an organism is called
- heterochrony.
 - developmental plasticity.
 - adaptation.
 - modularity.
 - mutation.
8. Which of the following is true about temperature determination of sex in some reptiles?
- It ensures that males and females are produced at different seasons of the year.
 - It ensures that males are faster than females.
 - It acts through the inactivation of the male sex chromosome.
 - There is no evidence that it has evolved because of effects on reproductive success.
 - Temperature effects are due to modifications of concentrations and actions of sex steroids.
9. Which of the following examples of evolutionary change do *not* involve Hox genes?
- Difference in numbers of legs between bees and centipedes.
 - Difference in number of cervical vertebrae between a goose and a giraffe.
 - Loss of forelegs in snakes.
 - Loss of webbing in the feet of chickens.
 - Location of legs and antennae in *Drosophila*.
10. Parallel phenotypic evolution is likely to occur because
- closely related organisms typically face similar problems.
 - the conservation of regulatory genes during evolution means that similar traits are likely to evolve repeatedly.
 - many different phenotypes can be produced by a given genotype.
 - phenotypic plasticity, which generates parallel phenotypic evolution, is widespread.
 - evolutionary biologists have looked especially hard to find evidence of it.

FOR DISCUSSION

- What environmental influences on development would probably be missed if investigations were confined to unicellular organisms such as bacteria and single-celled eukaryotes?
- If evolutionary innovations can result from rather simple changes in the timing of expression of a few genes, why have such innovations arisen relatively infrequently during evolution?
- François Jacob stated that evolution was more like tinkering than engineering. Does the observation that developmental

genes have changed little over evolutionary time support his assertion? Why?

- Despite their major differences, plants and animals share many of the genes that regulate development. What are the implications of this observation for the ways in which humans can respond to the adverse effects of the many substances we release into the environment that cause developmental abnormalities in plants and animals? What kinds of substances are most likely to have such effects? Why?

ADDITIONAL INVESTIGATION

Figure 20.7 describes an experiment in which the protein Gremlin, which inhibits expression of the *BMP4* gene, was introduced into the foot of a developing chicken. What results would you expect from introducing Gremlin into other parts of

a developing chicken? Why? Into what other body parts would it be most informative to introduce Gremlin? If you were particularly interested in parallel phenotypic evolution, what other organisms might you use in these experiments?

21

Evidence and Mechanisms of Evolution

Evolutionary theory leads to better flu vaccines

In November 11, 1918, an armistice agreement signed in France signaled the end of the First World War. But the death toll from four years of war was soon surpassed by the casualties of a massive influenza epidemic that began in the spring of 1918 among soldiers in a U.S. Army barracks. Over the next year and a half, this particular strain of flu virus spread across the globe in a true pandemic that killed more than 50 million people worldwide—more than twice the number of WWI-related combat deaths.

The 1918–1919 pandemic was also noteworthy because the death rate among young adults—who are usually less likely to die from influenza than are the elderly or the very young—was 20 times higher than in influenza epidemics before or since. Why was that particular flu virus so deadly, especially to typically hardy individuals?

Influenza viruses evolve constantly and rapidly. The 1918 strain triggered an especially intense reaction in the human immune system. This overreaction (called a “cytokine storm”) meant that people with strong immune systems were likely to be more severely affected. Usually, however, our immune system helps fight flu viruses, and this immune response is the basis of *vaccination* (see Chapter 42). The first flu vaccines were offered in 1945, and since then immunization programs have helped keep the number and severity of outbreaks in check. But flu viruses evolve so rapidly that last year’s vaccine is not effective against this year’s virus.

Although many of the influenza strains circulating in a given season are closely related, they are not identical. *Genetic variation* ensures that there are always different strains, and these different strains compete with one another. The strains that are best able to escape detection by the immune systems of their hosts are most likely to spread, and thus have an advantage over other strains. But the immune system responds to counteract the virus, and last year’s virus loses its advantage. If flu viruses did not evolve, we would become resistant and annual vaccination would be unnecessary.

Our immune system recognizes an influenza infection by detecting a protein called hemagglutinin on the viral surface. New mutations arise rapidly, and genetic sequence changes in the viral genome sometimes result in variation in the structure of hemagglutinin. Variants with altered hemagglutinin structure



Deadly Epidemic So many people were incapacitated during the flu epidemic of 1918–1919 that temporary hospitals had to be set up. Here the beds of flu-stricken patients cover the floor of the Dartmouth College gymnasium during the peak of the epidemic in the United States.



Reconstructing a Deadly Virus Terrence Tumpey, of the Centers for Disease Control and Prevention, reconstructed the 1918 influenza virus in his lab to identify the characteristics that made it so deadly.

are more likely to escape detection by our immune system, and thus more likely to survive and replicate. The term *positive selection* describes the evolution of favored changes like those in influenza hemagglutinin protein.

In developing flu vaccines, biologists examine the hemagglutinin gene, particularly at certain sites known to be under positive selection for change. Viruses with the greatest number of changes at these sites are the ones most likely to cause next year's flu epidemic, and therefore are the best targets for new vaccines. Understanding evolutionary theory thus helps us determine the causes and find solutions for a potentially deadly disease.

IN THIS CHAPTER we will examine the factual basis of evolution and consider some of the mechanisms that result in evolutionary change. We will see how Charles Darwin developed his ideas on one such mechanism (natural selection). We will discuss the genetic basis of evolution and show how genetic variation in populations is measured. Throughout the chapter, we will discuss ways that evolution can be applied to practical problems and how it helps us understand the diversity of life.

CHAPTER OUTLINE

- 21.1** What Facts Form the Basis of Our Understanding of Evolution?
- 21.2** What Are the Mechanisms of Evolutionary Change?
- 21.3** How Does Natural Selection Result in Evolution?
- 21.4** How Is Genetic Variation Maintained within Populations?
- 21.5** What Are the Constraints on Evolution?

21.1 What Facts Form the Basis of Our Understanding of Evolution?

The living world is constantly changing. Biologists observe many of these changes directly, both in laboratory experiments and in natural populations. Many other changes are recorded in the fossil record of life. We can measure the rate at which new mutations arise, observe the spread of new genetic variants through a population, and see the effects of genetic change on the form and function of organisms. In other words, evolution is a fact that we can observe directly. Biologists also have accumulated a large body of evidence about *how* evolutionary changes occur, and about *what* evolutionary changes have occurred in the past. The understanding and application of the mechanisms of evolutionary change to biological problems is known as evolutionary theory.

Evolutionary theory has many useful applications, such as the development of influenza vaccines described in the opening of this chapter. We use evolution to study, understand, and treat diseases; to develop better agricultural crops and practices; and to develop industrial processes that produce new molecules with useful properties. Knowledge of evolutionary principles has helped biologists understand how life diversified and how species interact. It also allows us to make predictions about the biological world.

In everyday speech, people tend to use the word "theory" to mean an untested hypothesis, or even a guess. But the term "evolutionary theory" does not refer to any single hypothesis, and it certainly is not guesswork. As used in science, "theory" refers to the entire body of work on the understanding and application of a field of knowledge. When we refer to "gravitational theory," we are not implying that gravity is an untested idea. No one doubts that gravity exists—we can see its effects all around us. Instead, we are referring to our understanding of the mechanisms that result in gravitational pull, and the use of that understanding to make predictions about the interactions of physical objects. Drop this book, and it will fall at a predictable rate, according to gravitational theory.

In a similar manner, when we refer to evolutionary theory, we are referring to our understanding of the mechanisms that result in biological changes in populations over time, and the use of that understanding to interpret changes and interactions of biological organisms. That biological populations evolve through time is not disputed by biologists. We can, and do, observe evolutionary change on a regular basis. We can directly observe the evolution of influenza viruses, but it is evolution-

ary theory that allows us to apply that information to the problem of developing more effective vaccines.

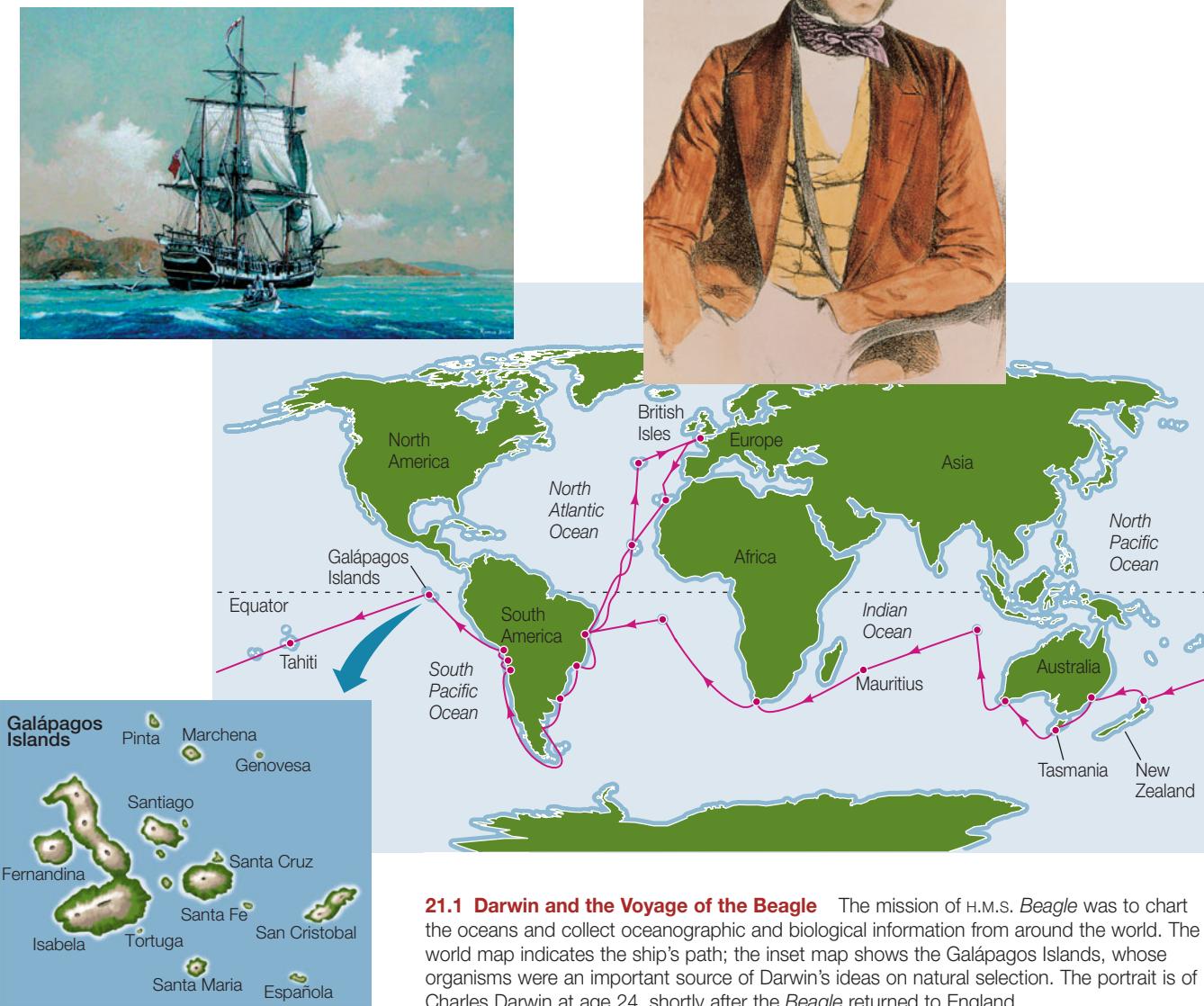
Several mechanisms of evolutionary change are recognized, and the scientific community is continually expanding its understanding of how and when these mechanisms apply to particular biological problems. Studying the mechanisms of evolution and their innumerable applications constitutes the active and exciting field of evolutionary theory.

Charles Darwin articulated the principle of natural selection

Today a rich array of geological, morphological, and molecular data support and enhance the factual basis of evolution. In the 1820s, however, it was not yet evident to the young Charles Darwin (or almost anyone else) that life had evolved. Darwin was passionately interested in both geology and natural history—the scientific study of how different organisms function and carry out their lives in nature. Despite these interests, he planned (at his father's behest) to become a doctor. But surgery conducted

without anesthesia nauseated Darwin, and he gave up medicine to study at Cambridge University for a career as a clergyman in the Church of England. Always more interested in science than in theology, he gravitated toward scientists on the faculty, especially the botanist John Henslow. In 1831, Henslow recommended Darwin for a position on the H.M.S. *Beagle*, which was preparing for a survey voyage around the world (Figure 21.1).

Whenever possible during the 5-year voyage, Darwin (who was often seasick) went ashore to study rocks and to observe and collect plants and animals. He noticed striking differences between the species he saw in South America and those from Europe. He observed that the species of the temperate regions of South America (Argentina and Chile) were more similar to those of tropical South America (Brazil) than they were to temperate European species. When he explored the islands of the



21.1 Darwin and the Voyage of the Beagle The mission of H.M.S. *Beagle* was to chart the oceans and collect oceanographic and biological information from around the world. The world map indicates the ship's path; the inset map shows the Galápagos Islands, whose organisms were an important source of Darwin's ideas on natural selection. The portrait is of Charles Darwin at age 24, shortly after the *Beagle* returned to England.

Galápagos Archipelago west of Ecuador, he noted that most of its animal species were found nowhere else, although they were similar to animals found on the mainland of South America. Darwin also observed that the animals of the Galápagos differed from island to island. He postulated that some animals had come to the archipelago from mainland South America and had subsequently undergone different changes on each of the islands. He then wondered what might account for these changes.

When he returned to England in 1836, Darwin continued to ponder his observations. Within a decade he had developed the major elements of an explanatory theory for evolutionary change based on three major propositions:

- Species are not immutable; they change over time.
- Divergent species share a common ancestor.
- The mechanism that produces changes in species is **natural selection**, or the differential survival and reproduction of individuals in a population based on variation in their traits.

The revolutionary assertions in Darwin's first two propositions were that evolution is a historical fact that can be demonstrated to have taken place, and that species are related to one another through common descent. In 1844, he wrote a long essay on natural selection, which he described as the mechanism of evolution (his third proposition), but despite urging from colleagues, he was reluctant to publish it, preferring to assemble more evidence first.

Darwin's hand was forced in 1858 when he received a letter and manuscript from another traveling naturalist, Alfred Russel Wallace, who was studying the biota of the Malay Archipelago. Wallace asked Darwin to evaluate his manuscript, which included an explanation of natural selection almost identical to Darwin's. Darwin was at first dismayed, believing Wallace had preempted his idea. But parts of Darwin's 1844 essay, together with Wallace's manuscript, were presented to the Linnaean Society of London on July 1, 1858, thereby crediting both men for the idea of natural selection. Darwin then worked quickly to finish his own book, *The Origin of Species*, which was published the next year.

Although Darwin and Wallace independently articulated the concept of natural selection, Darwin developed his ideas first.

Furthermore, *The Origin of Species* provided exhaustive evidence from many fields to support both natural selection and evolution itself. Thus both concepts are more closely associated with Darwin than with Wallace.

The facts that Darwin used to conceive and develop his explanation of evolution by natural selection were familiar to most contemporary biologists. His insight was to perceive the significance of relationships among these facts. Both Darwin and Wallace were influenced by the ideas of the economist Thomas Malthus, who in 1838 published *An Essay on the Principle of Population*. Malthus argued that because the rate of human population growth is greater than the rate of increase in food production, unchecked growth inevitably leads to famine. Darwin saw parallels throughout nature. He recognized that populations of all species have the potential to rapidly increase in number. To illustrate this point, he used the following example:

Suppose...there are eight pairs of birds, and that only four pairs of them annually...rear only four young, and that these go on rearing their young at the same rate, then at the end of seven years...there will be 2048 birds instead of the original sixteen.

Such increases are rarely seen in nature, though. Darwin therefore reasoned that death rates in nature must also be high. If they weren't, even the most slowly reproducing species would quickly reach enormous population sizes.

Darwin also observed that although offspring tend to resemble their parents, the offspring of most organisms are not identical to one another or to their parents. He suggested that slight variations among individuals affect the chance that a given individual will survive and reproduce. Darwin called this differential survival and reproduction of individuals natural selection.

Darwin may have used the words "natural selection" because he was familiar with the **artificial selection** of strains with certain desirable traits by animal and plant breeders. Many of Darwin's observations on the nature of variation came from domesticated plants and animals. Darwin was a pigeon breeder, and he knew firsthand the astonishing diversity in color, size, form, and behavior that breeders could achieve (Figure 21.2). He recognized



21.2 Artificial Selection Charles Darwin raised pigeons as a hobby, and he noted similar forces at work in artificial and natural selection. The "fancy" pigeons shown here represent three of the more than 300 varieties derived from the wild rock dove *Columba livia* (at left) by artificial selection for character traits such as color and feather distribution.

close parallels between selection by breeders and selection in nature. As he argued in *The Origin of Species*,

How can it be doubted, from the struggle each individual has to obtain subsistence, that any minute variation in structure, habits or instincts, adapting that individual better to the new conditions, would tell upon its vigour and health? In the struggle it would have a better chance of surviving; and those of its offspring which inherited the variation, be it ever so slight, would have a better chance.

That statement, written more than 150 years ago, still stands as a good expression of the process of evolution by natural selection.

It is important to remember that, as Darwin clearly understood, *individuals do not evolve; populations do*. A **population** is a group of individuals of a single species that live and interbreed in a particular geographic area at the same time. A major consequence of the evolution of populations is that their members become adapted to the environments in which they live. But what do biologists mean when they say that an organism is adapted to its environment?

yourBioPortal.com

GO TO Animated Tutorial 21.1 • Natural Selection

Adaptation has two meanings

In evolutionary biology, **adaptation** refers both to the *processes* by which characteristics that appear to be useful to their bearers evolve—that is, the evolutionary mechanisms that produce them—and to the *characteristics* themselves. With respect to the latter, an adaptation is a phenotypic characteristic that has made it more likely for an organism to survive and reproduce.

Biologists regard an organism as being adapted to a particular environment when they can demonstrate that a slightly different organism reproduces and survives less well in that environment. To understand adaptation, biologists compare the performance of individuals that differ in their traits. For example, biologists can assess the adaptive role of changes to the hemagglutinin protein of influenza viruses, as described in the opening of this chapter. By comparing the survival and proliferation rates of influenza viruses that have different hemagglutinin gene sequences, biologists can study adaptation of the viruses through time.

When Darwin proposed his ideas on evolution by natural selection, he could point to many examples of evolutionary mechanisms operating in nature, but none were supported by experiments. Since then, biologists have conducted thousands of observational and experimental studies that have confirmed the important role of natural selection as a mechanism of evolution. Biologists have also documented changes over time in the genetic composition and morphology of many populations, and our understanding of the mechanisms of inheritance has improved enormously since Darwin's time.

Population genetics provides an underpinning for Darwin's theory

Darwin had no knowledge of the mechanisms of genetic transmission, and his speculations on the topic proved to be incor-

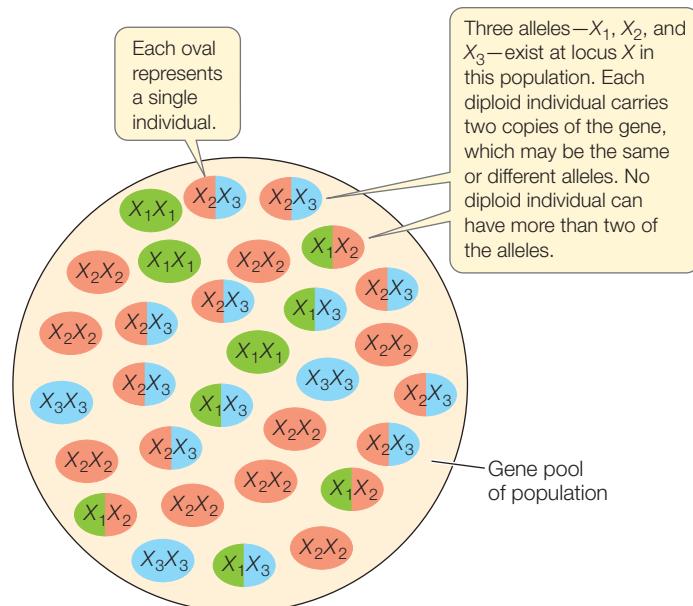
rect. Biologists did not have a good understanding of the genetic details of how natural selection works until the field of transmission genetics was established in the early 1900s. At that time, the rediscovery of Gregor Mendel's publications (see Section 12.1) paved the way for the development in the 1930s and 1940s of the field of *population genetics*. As the principles of evolution were integrated with the principles of modern genetics during this period, a new understanding of evolutionary biology—known as the *Modern Synthesis*—emerged. This was when biologists began to study mechanistic aspects of evolution as well as the broad evolutionary patterns that were so evident in nature.

For a population to evolve, its members must possess heritable genetic variation, which is the raw material on which mechanisms of evolution act. In everyday life, we do not directly observe the genetic compositions of organisms. What we see are *phenotypes*, the physical expressions of organisms' genes (including interactions among genes). The features of a phenotype are its *characters*—eye color, for example. The specific form of a character, such as brown eyes, is a *trait*. A **heritable trait** is a characteristic that is at least partly determined by the organism's genes. The genetic constitution that governs a character is called its *genotype*. A *population evolves when individuals with different genotypes survive or reproduce at different rates*.

The field of population genetics has three main goals:

- To explain the patterns and organization of genetic variation
- To explain the origin and maintenance of genetic variation
- To understand the mechanisms that cause changes in allele frequencies in populations

The perspective of population genetics complements the insights into evolutionary processes provided by developmental biology, as described in Chapter 20.



21.3 A Gene Pool A gene pool is the sum of all the alleles found in a population, or for a particular locus. This figure shows the gene pool for one locus, X. The allele frequencies are 0.20 for X_1 , 0.50 for X_2 , and 0.30 for X_3 .

21.4 Many Vegetables from One Species All of the crop plants shown here derive from a single wild mustard species. European agriculturalists produced these crop species by selecting and breeding plants with unusually large buds, stems, leaves, or flowers. The results substantiate the vast amount of variation present in a gene pool.

Different forms of a gene, known as **alleles**, may exist at a particular locus. At any particular locus, a single individual has only some of the alleles found in the population to which it belongs (Figure 21.3). The sum of all copies of all alleles at all loci found in a population constitutes its **gene pool**. (We can also refer to the “gene pool” for a particular locus or loci.) The gene pool is the source of the genetic variation that produces the phenotypic traits on which natural selection acts. To understand evolution and the role of natural selection, we need to know how much genetic variation populations have, what the sources of that genetic variation are, and how genetic variation changes in populations over space and time.

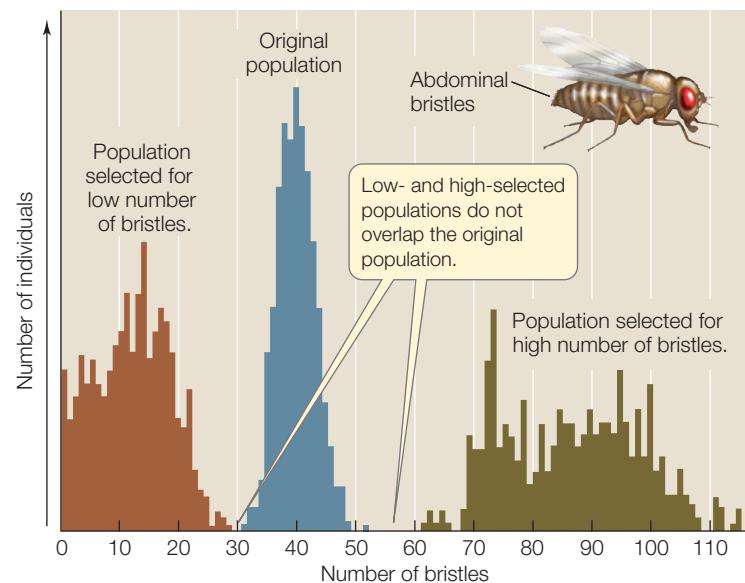
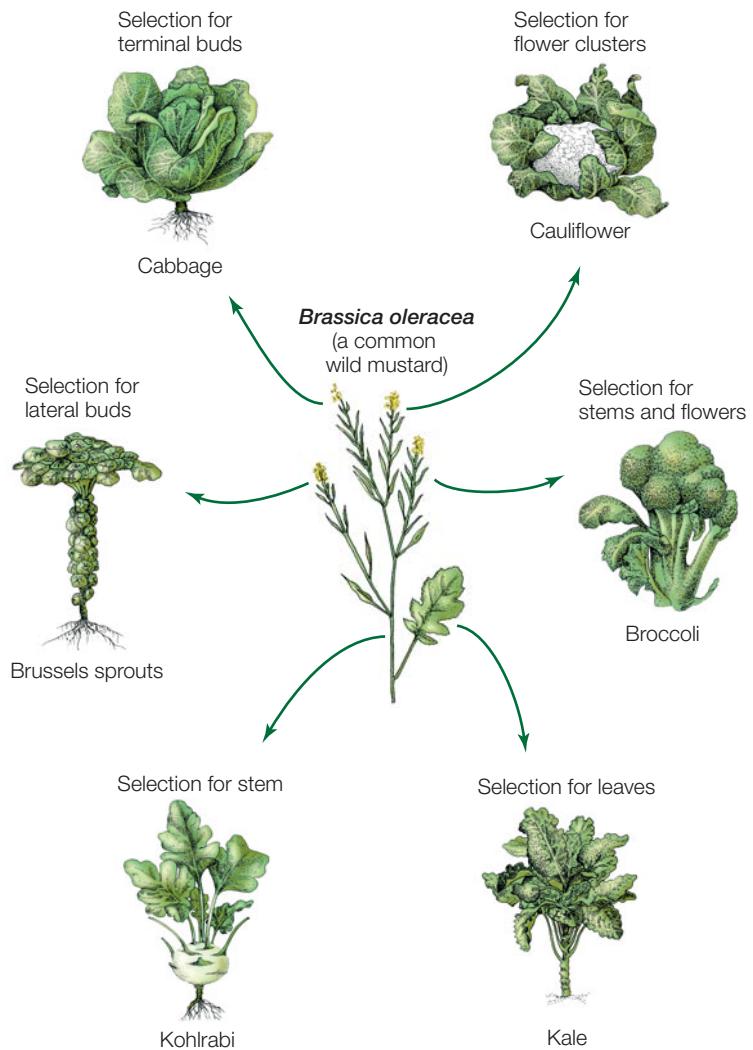
Most populations are genetically variable

Nearly all populations have genetic variation for many characters. Artificial selection on different characters in a single European species of wild mustard produced many important crop plants (Figure 21.4). Agriculturalists could achieve these results because the original mustard population had genetic variation for the characters of interest.

Laboratory experiments also demonstrate the existence of considerable genetic variation in populations. In one such experiment, investigators attempted to breed populations of the fruit fly *Drosophila melanogaster* with high or low numbers of bristles on their abdomens from an initial population with intermediate numbers of bristles. After 35 generations, all flies in both the high- and low-bristle lineages had bristle numbers that fell well outside the range found in the original population (Figure 21.5). Thus there must have been considerable genetic variation in the original fruit fly population on which selection could act.

Studying the genetic basis of natural selection is difficult because genotypes alone do not determine all phenotypes. With dominance, for example, a particular phenotype can be produced by more than one genotype (e.g., *AA* and *Aa* individuals may be phenotypically identical). Also, as we describe in Section 20.5, a given genotype can produce different phenotypes, depending on the environment encountered during development. For example, the cells of all the leaves on a tree or shrub are usually genetically identical, yet leaves of the same plant often differ in shape and size depending, for example, on the amount of ambient light they receive.

21.5 Artificial Selection Reveals Genetic Variation In experiments subjecting *Drosophila melanogaster* to artificial selection for bristle number, this trait evolved rapidly. The graphs show the number of flies with different numbers of bristles after 35 generations of artificial selection, which clearly diverged from the bristle numbers present in the original population (the blue bars in the center).



Evolutionary change can be measured by allele and genotype frequencies

Allele frequencies are usually estimated in locally interbreeding groups, called Mendelian populations, within a geographic population of a species. To measure allele frequencies in a Mendelian population precisely, we would need to count every allele at every locus in every individual in the population. By doing so, we could determine the frequencies of all alleles in the population. The word *frequency* in this case refers to an allele's proportion in the gene pool at a particular locus.

Fortunately, we do not need to make complete measurements, because we can reliably estimate allele frequencies for a given locus by counting alleles in a sample of individuals from the population. The sum of all allele frequencies at a locus is equal to 1, so measures of allele frequency range from 0 to 1.

An allele's frequency is calculated using the following formula:

$$p = \frac{\text{number of copies of the allele in the population}}{\text{sum of alleles in the population}}$$

If only two alleles (we'll call them *A* and *a*) for a given locus are found among the members of a diploid population, they may combine to form three different genotypes: *AA*, *Aa*, and *aa*. Such a population is said to be *polymorphic* at that locus, since there is more than one allele. Using the formula above, we can calculate the relative frequencies of alleles *A* and *a* in a population of *N* individuals as follows:

- Let N_{AA} be the number of individuals that are homozygous for the *A* allele (*AA*).
- Let N_{Aa} be the number that are heterozygous (*Aa*).
- Let N_{aa} be the number that are homozygous for the *a* allele (*aa*).

Note that $N_{AA} + N_{Aa} + N_{aa} = N$, the total number of individuals in the population, and that the total number of copies of both alleles present in the population is $2N$, because each individual is diploid. Each *AA* individual has two copies of the *A* allele, and

each *Aa* individual has one copy of the *A* allele. Therefore, the total number of *A* alleles in the population is $2N_{AA} + N_{Aa}$. Similarly, the total number of *a* alleles in the population is $2N_{aa} + N_{Aa}$.

If *p* represents the frequency of *A*, and *q* represents the frequency of *a*, then

$$p = \frac{2N_{AA} + N_{Aa}}{2N}$$

and

$$q = \frac{2N_{aa} + N_{Aa}}{2N}$$

Figure 21.6 shows how these formulas can be used to calculate allele frequencies in two hypothetical populations, each containing 200 diploid individuals. Population 1 has mostly homozygotes (90 *AA*, 40 *Aa*, and 70 *aa*), whereas population 2 has mostly heterozygotes (45 *AA*, 130 *Aa*, and 25 *aa*).

The calculations in Figure 21.6 demonstrate two important points. First, notice that for each population, $p + q = 1$, which means that $q = 1 - p$. So when there are only two alleles at a given locus in a population, we can calculate the frequency of one allele and then obtain the second allele's frequency by subtraction. If there is only one allele at a given locus in a population, its frequency is 1: the population is then *monomorphic* at that locus, and the allele is said to be *fixed*.

The second thing to notice is that population 1 (consisting mostly of homozygotes) and population 2 (consisting mostly of heterozygotes) have the same allele frequencies for *A* and *a*. Thus they have the same gene pool for this locus. Because the alleles in the gene pool are distributed differently among individuals, however, the *genotype frequencies* of the two populations differ. Genotype frequencies are calculated as the number of individuals that have a given genotype divided by the total number of individuals in the population. Using the numbers in Figure 21.6, the genotype frequencies in population 1 would be 0.45 *AA*, 0.20 *Aa*, and 0.35 *aa*.

The frequencies of different alleles at each locus and the frequencies of different genotypes in a Mendelian population de-

TOOLS FOR INVESTIGATING LIFE

21.6 Calculating Allele Frequencies

Allele frequencies for any gene pool can be calculated using the equations in panel 1. When these equations are applied to the populations in panel 2, we find that the *frequencies* of alleles *A* and *a* in the two populations are the same, but the alleles are distributed differently between heterozygous and homozygous genotypes.

1 Determine the allele frequencies in the population.

In any population:

$$\text{Frequency of allele } A = p = \frac{2N_{AA} + N_{Aa}}{2N} \quad \text{Frequency of allele } a = q = \frac{2N_{aa} + N_{Aa}}{2N}$$

where *N* is the total number of individuals in the population.

2 Compute allele frequencies for different populations.

For population 1
(mostly homozygotes):

$$N_{AA} = 90, N_{Aa} = 40, \text{ and } N_{aa} = 70 \quad N_{AA} = 45, N_{Aa} = 130, \text{ and } N_{aa} = 25$$

so

$$p = \frac{180 + 40}{400} = 0.55$$

$$q = \frac{140 + 40}{400} = 0.45$$

For population 2
(mostly heterozygotes):

$$p = \frac{90 + 130}{400} = 0.55$$

$$q = \frac{50 + 130}{400} = 0.45$$

scribe that population's **genetic structure**. Allele frequencies measure the amount of genetic variation in a population; genotype frequencies show how a population's genetic variation is distributed among its members. Other measures, such as the proportion of polymorphic loci, are also used to measure variation in populations. With these measurements, it becomes possible to consider how the genetic structure of a population changes or remains the same over generations—that is, to measure evolutionary change.

The genetic structure of a population changes over time, unless certain restrictive conditions exist

In 1908, the British mathematician Godfrey Hardy and the German physician Wilhelm Weinberg independently deduced the conditions that must prevail if the genetic structure of a population is to remain the same over time. If the conditions they identified do not exist, then evolution will occur. The resulting principle, known as **Hardy–Weinberg equilibrium**, is a cornerstone of population genetics. Hardy–Weinberg equilibrium describes a model in which allele frequencies do not change across generations and genotype frequencies can be predicted from allele frequencies (Figure 21.7). The principles of Hardy–Weinberg equilibrium apply only to sexually reproducing organisms. Several conditions must be met for a population to be at Hardy–Weinberg equilibrium:

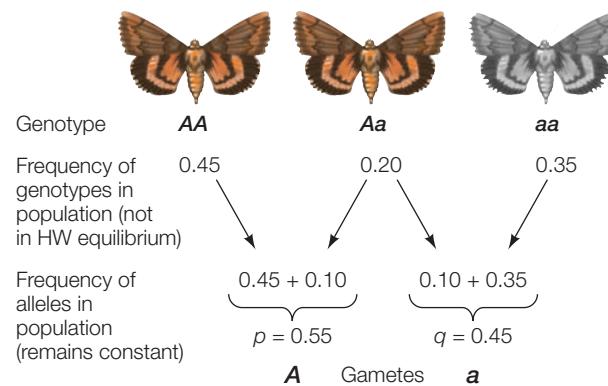
- **Mating is random.** Individuals do not preferentially choose mates with certain genotypes.
- **Population size is infinite.** The larger a population, the smaller will be the effect of **genetic drift**—random (chance) fluctuations in allele frequencies from one generation to another.
- **There is no gene flow** (movement of individuals into or out of the population, or reproductive contact with other populations).
- **There is no mutation.** There is no change to alleles in the population, and no new alleles are added to change the gene pool.
- **Selection does not affect the survival of particular genotypes.** There is no differential survival of individuals with different genotypes.

If these “ideal” conditions hold, two major consequences follow. First, the frequencies of alleles at a locus remain constant from generation to generation. Second, following one generation of random mating, the genotype frequencies occur in the following proportions:

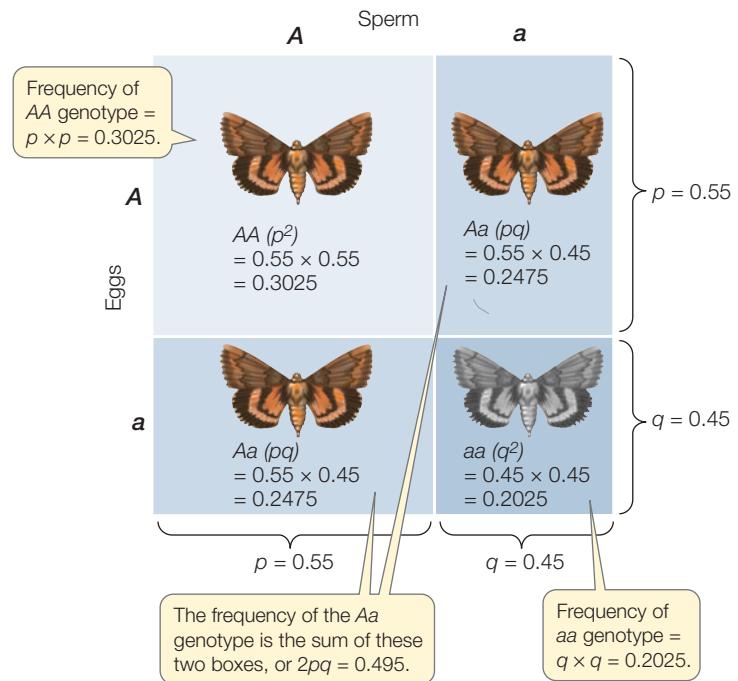
Genotype	AA	Aa	aa
Frequency	p^2	$2pq$	q^2

Consider a population that is *not* in Hardy–Weinberg equilibrium, such as generation I of Figure 21.7. This could occur, for example, if the initial population is founded by migrants from several other populations, thus violating the Hardy–Weinberg assumption of no gene flow. In this example, “generation I” has more homozygous individuals and fewer heterozygous individuals than would be expected under Hardy–Weinberg equilibrium (a condition known as *heterozygote deficiency*).

Generation I (Founder population)



Generation II (Hardy–Weinberg equilibrium restored)



21.7 One Generation of Random Mating Restores Hardy–Weinberg Equilibrium

Generation I of this example population is founded by migrants from several source populations, and so is not initially in Hardy–Weinberg equilibrium. After one generation of random mating, the allele frequencies are unchanged and the genotype frequencies return to Hardy–Weinberg expectations. The length of the sides of each rectangle are proportional to the allele frequencies in the population; the areas of the rectangles are proportional to the genotype frequencies.

Even with a starting population that is not in Hardy–Weinberg equilibrium, we would predict that after a single generation of random mating, and without violating the other Hardy–Weinberg assumptions, the *allele frequencies* will remain unchanged but the *genotype frequencies* will return to Hardy–Weinberg expectations. Let's explore why this is true.

In generation I of Figure 21.7, the frequency of the A allele (p) is 0.55. Because we assume that individuals select mates at random, without regard to their genotype, gametes carrying A

or a combine at random—that is, as predicted by the allele frequencies of p and q . Thus in this example, the probability that a particular sperm or egg will bear an A allele is 0.55. In other words, 55 out of 100 randomly sampled sperm or eggs will bear an A allele. Because $q = 1 - p$, the probability that a sperm or egg will bear an a allele is $1 - 0.55 = 0.45$. (You may wish to review the discussion of probability in Section 12.1.)

To obtain the probability of two A -bearing gametes coming together at fertilization, we multiply the two independent probabilities of their occurring separately:

$$p \times p = p^2 = (0.55)^2 = 0.3025$$

Therefore, 0.3025, or 30.25 percent, of the offspring in generation II will have homozygous genotype AA . Similarly, the probability of bringing together two a -bearing gametes is

$$q \times q = q^2 = (0.45)^2 = 0.2025$$

Thus 20.25 percent of generation II will have the aa genotype.

There are two ways of producing a heterozygote: an A sperm may combine with an a egg, the probability of which is $p \times q$; or an a sperm may combine with an A egg, the probability of which is $q \times p$. Consequently, the overall probability of obtaining a heterozygote is $2pq$, or 0.495. The frequencies of the AA , Aa , and aa genotypes in generation II of Figure 21.7 are now at Hardy–Weinberg expectations, and the frequencies of the two alleles (p and q) have not changed from generation I.

Under the assumptions of Hardy–Weinberg equilibrium, allele frequencies p and q remain constant for each generation. If Hardy–Weinberg assumptions are violated and the genotype frequencies in the parental generation are altered (say, by the loss of a large number of AA individuals from the population), then the allele frequencies in the next generation would be altered. However, based on the new allele frequencies, another generation of random mating is sufficient to restore the genotype frequencies to Hardy–Weinberg equilibrium.

yourBioPortal.com

GO TO Animated Tutorial 21.2 • Hardy–Weinberg Equilibrium

Deviations from Hardy–Weinberg equilibrium show that evolution is occurring

You probably have realized that populations in nature never meet the stringent conditions necessary to be at Hardy–Weinberg equilibrium. Why, then, is this model considered so important for the study of evolution? There are two reasons. First, the equation is often useful for predicting the approximate genotype frequencies of a population from its allele frequencies. Second—and crucially—the model describes the conditions required for there to be *no* evolution in a population.

Few if any of the Hardy–Weinberg model’s conditions are ever met completely in real populations, and allele frequencies in all populations do in fact change through time—that is, populations *do* evolve. The specific patterns of deviation from Hardy–Weinberg equilibrium can help us identify the various mechanisms of evolutionary change.

21.1 RECAP

Evolutionary change is directly observable in biological populations. Genetic variation is needed for evolutionary change to occur, and natural selection is one of the major mechanisms that acts on that variation and results in evolution.

- Can you articulate the principle of natural selection? The two meanings of adaptation? [See pp. 442–444](#)
- How does calculating allele frequencies allow us to measure evolutionary change? [See pp. 446–447 and Figure 21.6](#)
- Why is the concept of Hardy–Weinberg equilibrium important even though the assumptions on which it is based are never completely met in nature? [See pp. 447–448](#)

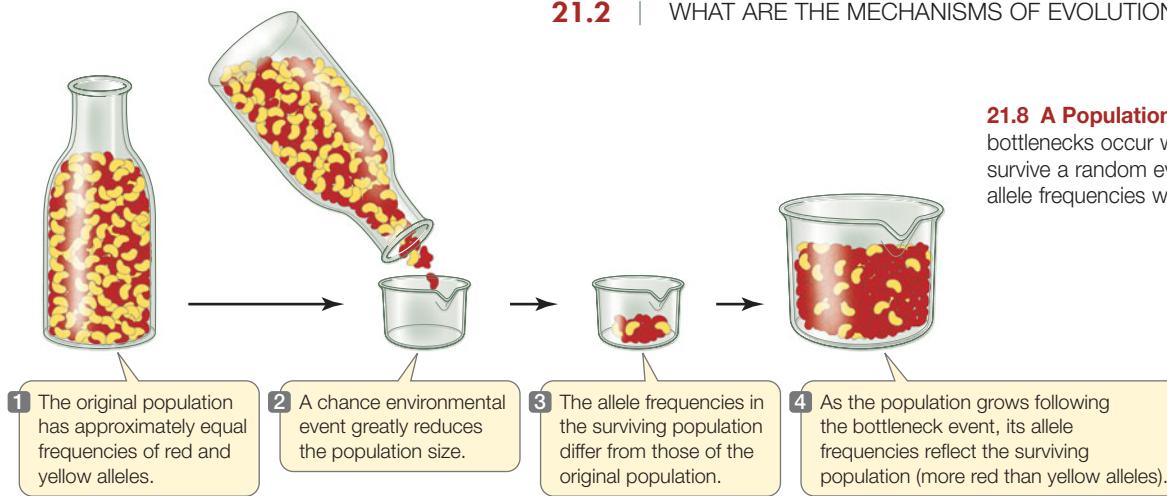
We have briefly outlined Charles Darwin’s vision of natural selection and adaptation and explained the mathematical basis of Hardy–Weinberg equilibrium and its importance for studying evolution. We’ll now examine some of the forces that cause populations to deviate from equilibrium—the mechanisms of evolutionary change.

21.2 What Are the Mechanisms of Evolutionary Change?

Evolutionary mechanisms are forces that change the genetic structure of a population. Hardy–Weinberg equilibrium is a null hypothesis that assumes those forces are absent. The known evolutionary mechanisms include mutation, gene flow, genetic drift, nonrandom mating, and selection—each of which contradicts one of the five basic assumptions of Hardy–Weinberg equilibrium. We have already discussed Darwin’s principal explanation for evolution, namely natural selection. Although natural selection is in many cases an important component of evolution, even Darwin recognized that it was not the only mechanism of evolution, and many additional evolutionary forces have been discovered since Darwin’s time. Here we discuss some of the other mechanisms that result in evolution.

Mutations generate genetic variation

The origin of genetic variation is mutation. A **mutation**, as Section 14.6 describes, is any change in the nucleotide sequences of an organism’s DNA. The process of DNA replication is not perfect, and changes appear almost every time a genome is replicated. Mutations occur randomly with respect to an organism’s adaptive needs; it is selection acting on this random variation that results in adaptation. Most mutations are either harmful to their bearers or neutral. A few are beneficial, however, and previously harmful or neutral alleles may become advantageous if conditions change. In addition, mutations can restore to a population genetic variation that other evolutionary processes have



21.8 A Population Bottleneck Population bottlenecks occur when only a few individuals survive a random event, resulting in a shift in allele frequencies within the population.

removed. Thus mutations both create and help maintain genetic variation in populations.

Mutation rates can be high, as we saw with the influenza viruses described in the opening of this chapter, but in many organisms the mutation rate is very low (on the order of 10^{-8} to 10^{-9} changes per base pair of DNA per generation). Even low overall mutation rates, however, are sufficient to create considerable genetic variation, because each of a large number of genes may change, and populations often contain large numbers of individuals. For example, if the probability of a point mutation (an addition, deletion, or substitution of a single base) were 10^{-9} per base pair per generation, then each human gamete, the DNA of which contains 3×10^9 base pairs, would average three new point mutations ($3 \times 10^9 \times 10^{-9} = 3$)—and each zygote would carry an average of six new mutations. The current human population of about 7 billion people would be expected to carry about 42 billion new mutations that were not present one generation earlier. So even though the mutation rate in humans is quite low, human populations still contain enormous genetic variation on which selection can act.

One of the conditions for Hardy–Weinberg equilibrium is that there be no mutation. Although this condition is never strictly met, the rate at which mutations arise at a single locus is usually so low that mutations by themselves result in only small deviations from Hardy–Weinberg equilibrium. If large deviations are found, it is usually appropriate to dismiss mutation as the cause and to look for evidence of other evolutionary mechanisms acting on the population.

Gene flow may change allele frequencies

Few populations are completely isolated from other populations of the same species. Migration of individuals and movements of gametes between populations—a phenomenon called gene flow—can change allele frequencies in a population. If the arriving individuals survive and reproduce in their new location, they may add new alleles to the population’s gene pool, or they may change the frequencies of alleles already present if they come from a population with different allele frequencies. For a population to be at Hardy–Weinberg equilibrium, there must be no gene flow from populations with different allele frequencies.

Genetic drift may cause large changes in small populations

In small populations, genetic drift—random changes in allele frequencies from one generation to the next—may produce large changes in allele frequencies over time. Harmful alleles may increase in frequency, and rare advantageous alleles may be lost. Even in large populations, genetic drift can influence the frequencies of alleles that do not affect the survival and reproductive rates of their bearers.

As an example, suppose we cross $Aa \times Aa$ fruit flies to produce an F_1 population in which $p = q = 0.5$ and in which the genotype frequencies are 0.25 AA , 0.50 Aa , and 0.25 aa . If we randomly select 4 individuals (= 8 copies of the gene) from the F_1 population to produce the F_2 generation, the allele frequencies in this small sample population may differ markedly from $p = q = 0.5$. If, for example, we happen by chance to draw 2 AA homozygotes and 2 heterozygotes (Aa), the allele frequencies in the sample will be $p = 0.75$ (6 out of 8) and $q = 0.25$ (2 out of 8). If we replicate this experiment 1,000 times, one of the two alleles will be missing entirely from about 8 of the 1,000 sample populations.

The same principles operate when a population is reduced dramatically in size. Populations that are normally large may occasionally pass through a period in which only a small number of individuals survive, a situation known as a **population bottleneck**. During population bottlenecks, genetic variation can be reduced by genetic drift. This is illustrated in Figure 21.8, in which red and yellow beans represent two different alleles of a gene. Most of the “surviving” beans in the small sample taken from the original population are, just by chance, red, so the new population has a much higher frequency of red beans than the previous generation had. In a real population, the allele frequencies would be described as having “drifted.”

A population forced through a bottleneck is likely to lose much of its genetic variation. For example, when Europeans first arrived in North America, millions of greater prairie-chickens (*Tympanuchus cupido*) inhabited the prairies. As a result of hunting and habitat destruction by the new settlers, the Illinois population of this species plummeted from about 100 million birds in 1900 to fewer than 50 in the 1990s (Figure 21.9A). A comparison of DNA from birds collected in Illinois during the middle of the twentieth century with DNA from the surviving pop-

(A) *Tympanuchus cupido*(B) *Washingtonia filifera*

21.9 Species with Low Genetic Variation (A) Greater prairie-chickens in Illinois lost most of their genetic variation when the population crashed from millions to fewer than 50 individuals. (B) The California fan palm, whose range has been reduced to a small area of southern California and neighboring Mexico, has little genetic variation.

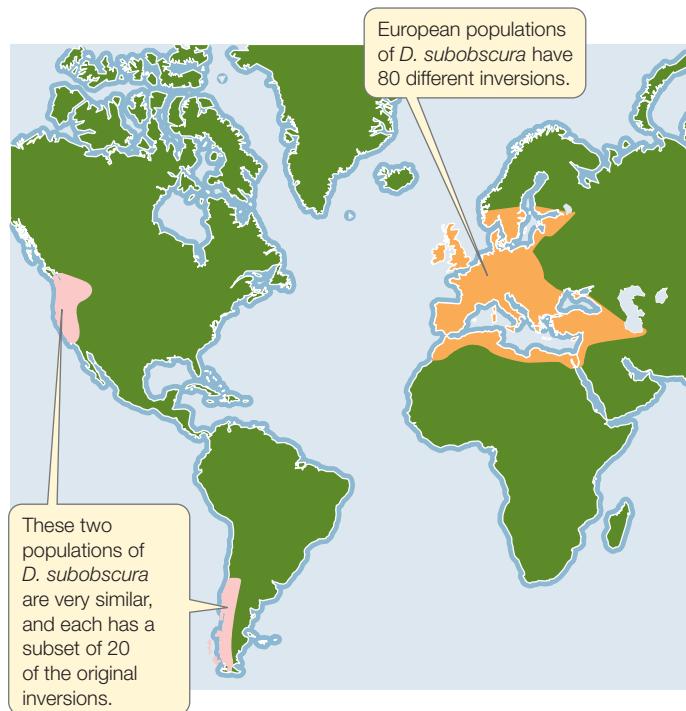
ulation in the 1990s showed that Illinois prairie-chickens have lost most of their genetic diversity. The remaining population is experiencing low reproductive success. Similarly, the California fan palm (*Washingtonia filifera*) was once widespread in California and Mexico; today it is restricted to a few oases in extreme southern California and adjacent Mexico (Figure 21.9B). The species has little genetic variation: an average individual is heterozygous at fewer than 1 percent of its loci.

Genetic drift can have similar effects when a few pioneering individuals colonize a new region. Because of its small size, the

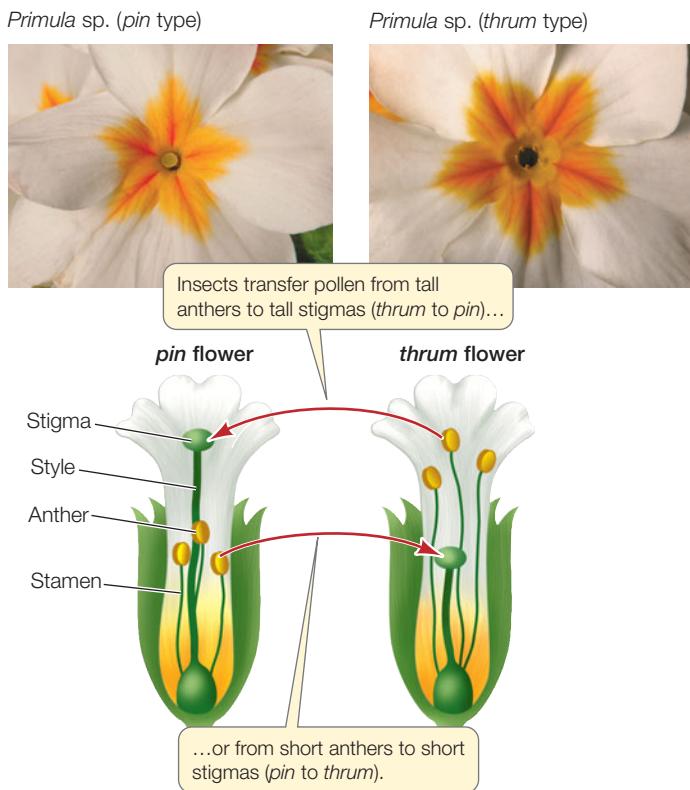
colonizing population is unlikely to have all the alleles found among members of its source population. The resulting change in genetic variation, called a **founder effect**, is equivalent to that in a large population reduced by a bottleneck. For example, the current population of the pitcher plant *Sarracenia purpurea* on a small island in central Ohio arose from a single individual that was planted there in 1912. Today the population has only one detectable polymorphic locus in its entire genome.

Scientists were given an opportunity to study the genetic composition of founding populations when *Drosophila subobscura*, a well-studied species of fruit fly native to Europe, was discovered near Puerto Montt, Chile (in 1978), and at Port Townsend, Washington (in 1982). The *D. subobscura* founders probably reached Chile from Europe on a ship, and a few flies carried north from Chile on another ship founded the North American population. In both South and North America, populations of the flies grew rapidly and expanded their ranges. Today in North America, *D. subobscura* ranges from British Columbia to central California. In Chile it has spread across 15 degrees of latitude (Figure 21.10).

European populations of *D. subobscura* have 80 different chromosomal inversions, but the North and South American populations have only a subset of 20 of these inversions—and they are the same 20 on both continents. North and South American populations also have lower allele diversity at certain enzyme-producing genes compared with European populations. Only those alleles that have a frequency higher than 0.10 in European populations are present in the Americas. Thus, as expected for a small founding population, only a small part of the total genetic variation found in Europe reached the Americas. Geneticists estimate that somewhere between 4 and 100 flies founded the North and South American populations.



21.10 A Founder Effect Populations of the fruit fly *Drosophila subobscura* in North and South America contain less genetic variation than the European populations from which they came, as measured by the number of chromosome inversions in each population. Within two decades of arriving in the Americas, *D. subobscura* populations had increased dramatically and spread widely in spite of their reduced genetic variation.



21.11 Flower Structure Fosters Nonrandom Mating Differing floral structure within the same plant species, as illustrated by this primrose, ensures that pollination usually occurs between individuals of different genotypes.

Nonrandom mating can change genotype frequencies

Mating patterns may alter genotype frequencies if individuals in a population do not choose mates at random. For example, if they mate preferentially with individuals of the same genotype, then homozygous genotypes will be overrepresented and heterozygous genotypes underrepresented relative to Hardy–Weinberg expectations. Alternatively, individuals may mate primarily or exclusively with individuals of different genotypes.

Nonrandom mating is seen in some plant species, such as primroses (genus *Primula*), in which individual plants bear flowers of only one of two different types. One type, known as *pin*, has a long style (the stalk that supports the stigma, where pollen is received) and short stamens (the stalks ending in anthers, where pollen is produced). The other type, known as *thrum*, has a short style and long stamens (Figure 21.11). In many species with this reciprocal arrangement, pollen from one flower type can fertilize only flowers of the other type. Pollen grains from *pin* and *thrum* flowers are deposited on different parts of the bodies of insects that visit the flowers. When the insects visit other flowers, pollen grains from *pin* flowers are most likely to come into contact with stigmas of *thrum* flowers, and vice versa.

Self-fertilization (*selfing*), another form of nonrandom mating, is common in many groups of organisms, especially plants. Selfing reduces the frequencies of heterozygous individuals

from Hardy–Weinberg equilibrium and increases the frequencies of homozygotes, but it does not change allele frequencies.

Sexual selection is a particularly important form of nonrandom mating that *does* change allele frequencies and also often results in the evolution of significant differences between males and females of a species. We will discuss this important evolutionary mechanism in detail in the next section.

21.2 RECAP

Evolutionary mechanisms are processes that change the genetic structure of a population. Known evolutionary mechanisms include mutation, gene flow, genetic drift, nonrandom mating, and selection.

- Why do mutations, by themselves, result in only small deviations from Hardy–Weinberg equilibrium?
See pp. 448–449
- Explain how genetic drift can cause large changes in small populations. **See pp. 449–450 and Figure 21.8**
- Why is it that some types of nonrandom mating alter genotype frequencies but not allele frequencies?
See p. 451

The evolutionary mechanisms discussed so far influence the frequencies of alleles and genotypes in populations. Although all of these processes influence the course of biological evolution, only natural selection results in adaptation. For adaptation to occur, individuals that differ in heritable traits must survive and reproduce with different degrees of success.

21.3 How Does Natural Selection Result in Evolution?

Although evolution is defined as changes in the gene frequencies of a population from one generation to the next, natural selection acts on the *phenotype*—the physical features expressed by an organism with a given genotype—rather than directly on the genotype. The reproductive contribution of a phenotype to subsequent generations relative to the contributions of other phenotypes is called its **fitness**.

Changes in reproductive rate do not necessarily change the genetic structure of a population. For example, if all individuals in a population experience the same increase in reproductive rate (during an environmentally favorable year, for instance), the genetic structure of the population will not change. Changes in numbers of offspring are responsible for increases and decreases in the *size* of a population, but only changes in the *relative success* of different phenotypes in a population lead to changes in allele frequencies from one generation to the next. The fitness of individuals of a particular phenotype is a function of the probability of those individuals surviving multiplied by the average number of offspring they produce over their lifetimes. In other words, the *fitness of a phenotype is determined by the relative rates of survival and reproduction of individuals with that phenotype*.

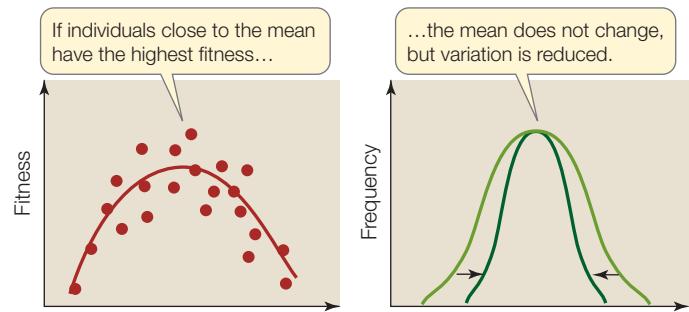
Natural selection can change or stabilize populations

To simplify our discussion until now, we have considered only characters influenced by alleles at a single locus. As we describe in Section 12.3, however, most characters are influenced by alleles at more than one locus. Such characters are likely to show quantitative rather than qualitative variation. For example, the distribution of body sizes of individuals in a population, a character that is influenced by genes at many loci as well as by the environment, is likely to resemble the bell-shaped curves shown in the right-hand column of Figure 21.12.

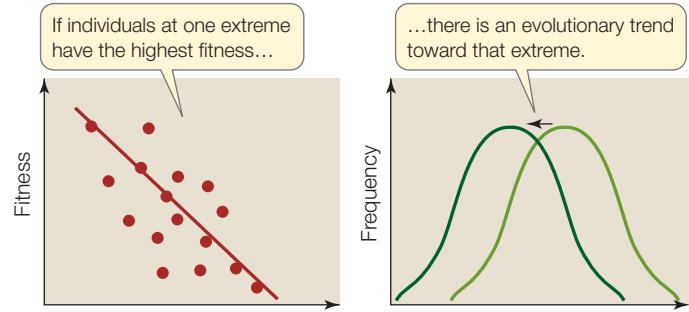
Natural selection can act on characters with quantitative variation in any one of several different ways, producing quite different results:

- **Stabilizing selection** preserves the average characteristics of a population by favoring average individuals.

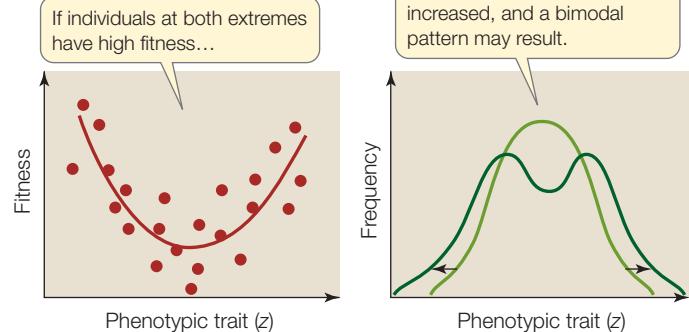
(A) Stabilizing selection



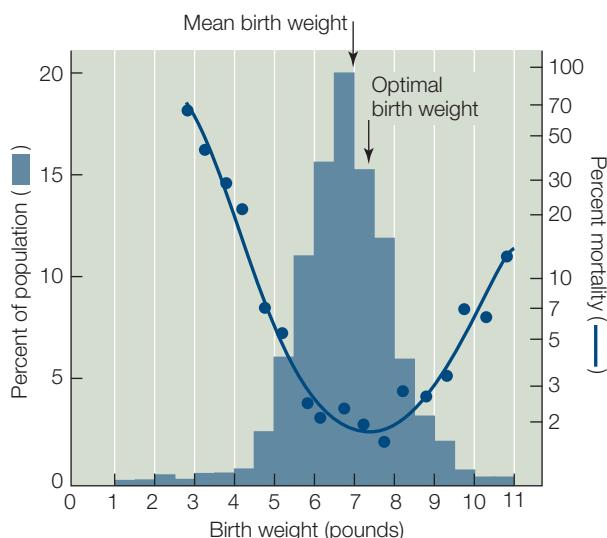
(B) Directional selection



(C) Disruptive selection



21.12 Natural Selection Can Operate in Several Ways The graphs in the left-hand column show the fitness of individuals with different phenotypes of the same trait. The graphs on the right show the distribution of the phenotypes in the population before (light green) and after (dark green) the influence of selection.



21.13 Human Birth Weight Is Influenced by Stabilizing Selection

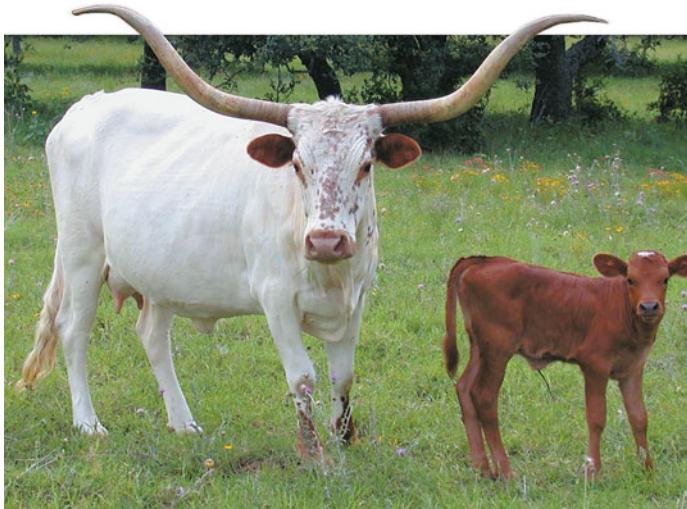
Babies that weigh more or less than average are more likely to die soon after birth than babies with weights close to the population mean.

- **Directional selection** changes the characteristics of a population by favoring individuals that vary in one direction from the mean of the population.
- **Disruptive selection** changes the characteristics of a population by favoring individuals that vary in both directions from the mean of the population.

STABILIZING SELECTION If the smallest and largest individuals in a population contribute fewer offspring to the next generation than do individuals closer to the average size, then stabilizing selection is operating on size (Figure 21.12A). Stabilizing selection reduces variation in populations, but it does not change the mean. Natural selection frequently acts in this way, countering increases in variation brought about by sexual recombination, mutation, or migration. Rates of evolution in many species are slow because natural selection is often stabilizing. Stabilizing selection operates, for example, on human birth weight. Babies born lighter or heavier than the population mean die at higher rates than babies whose weights are close to the mean (Figure 21.13). In discussions of specific genes, stabilizing selection is often called *purifying selection*, because there is selection against any deleterious mutations to the usual gene sequence.

DIRECTIONAL SELECTION Directional selection is operating when individuals at one extreme of a character distribution contribute more offspring to the next generation than other individuals do, shifting the average value of that character in the population toward that extreme. In the case of a single gene locus, directional selection may result in favoring a particular genetic variant (known as *positive selection* for that variant). By favoring one phenotype over another, directional selection results in an increase of the frequencies of alleles that produce the favored phenotype (as with the hemagglutinin gene of influenza in the opening of this chapter).

If directional selection operates over many generations, an *evolutionary trend* is seen in the population (Figure 21.12B). Directional evolutionary trends often continue for many genera-



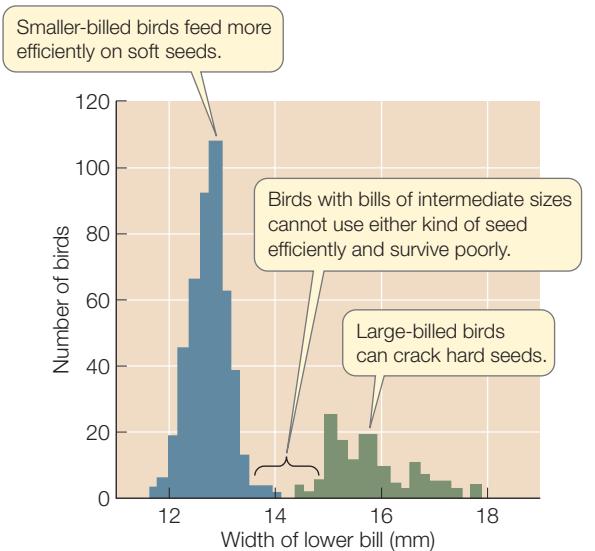
21.14 Texas Longhorns Are the Result of Directional Selection

Longer horns were advantageous for defending young calves from attacks by predators, so feral herds of Spanish cattle developed much longer horns between the early 1500s and the 1860s. The trend has been maintained in modern times by ranchers practicing artificial selection.

tions, but they can be reversed if the environment changes and different phenotypes are favored, or halted when an optimal phenotype is reached or when trade-offs oppose further change. The character then falls under stabilizing selection.

Many cases of directional selection have been observed directly, and long-term examples abound in the fossil record. The long horns of Texas Longhorn cattle (Figure 21.14) are an example of a trait that has evolved through directional selection. Texas Longhorns are descendants of cattle that Christopher Columbus brought to the New World. Columbus picked up a few cattle in the Canary Islands and brought them to the island of Hispaniola in 1493. The cattle quickly multiplied, and their descendants were taken to the mainland of Mexico. As the Spanish explored what would later become Texas and the southwestern United States, they brought some of these cattle with them, some of which escaped and formed feral herds. Populations of these feral cattle increased greatly over the next few hundred years, but there was heavy predation from bears, mountain lions, and wolves, especially on the young calves. Cows with longer horns were more successful in protecting their calves against attacks, and over the next few hundred years the average horn length of cattle in the feral herds increased considerably. In addition, the cattle evolved resistance to endemic diseases of the Southwest, as well as higher fecundity and longevity. Texas Longhorn cows often live and produce calves well into their twenties, or about twice as long as many breeds of cattle that have been artificially selected by humans for traits such as high fat content or high milk production (which are examples of artificial directional selection).

DISRUPTIVE SELECTION When disruptive selection operates, individuals at opposite extremes of a character distribution contribute more offspring to the next generation than do individuals close to the mean, which increases variation in the population (Figure 21.12C).



21.15 Disruptive Selection Results in a Bimodal Distribution

The bimodal distribution of bill sizes in the black-bellied seedcracker of West Africa is a result of disruptive selection, which favors individuals with larger and smaller bill sizes over individuals with intermediate-sized bills.

The strikingly bimodal (two-peaked) distribution of bill sizes in the black-bellied seedcracker (*Pyrenestes ostrinus*), a West African finch (Figure 21.15), illustrates how disruptive selection can influence populations in nature. The seeds of two types of sedges (marsh plants) are the most abundant food source for these finches during part of the year. Birds with large bills can readily crack the hard seeds of the sedge *Scleria verrucosa*. Birds with small bills can crack *S. verrucosa* seeds only with difficulty; however, they feed more efficiently on the soft seeds of *S. goossensis* than do birds with larger bills.

Young finches whose bills deviate markedly from the two predominant bill sizes do not survive as well as finches whose bills are close to one of the two sizes represented by the distribution peaks. Because there are few abundant food sources in the environment, and because the seeds of the two sedges do not overlap in hardness, birds with intermediate-sized bills are less efficient in using either one of the principal food sources. Disruptive selection therefore maintains a bimodal bill size distribution.

Sexual selection influences reproductive success

Sexual selection acts on characteristics that determine reproductive success. In *The Origin of Species*, Darwin devoted only a few pages to sexual selection, but in 1871 he wrote an entire book about it: *The Descent of Man, and Selection in Relation to Sex*. Sexual selection was Darwin's explanation for the evolution of conspicuous characters that would appear to inhibit survival, such as bright colors, long tails, and elaborate courtship displays in males of many species. He hypothesized that these features either improved the ability of their bearers to compete for access to mates (*intrasexual selection*) or made their bearers more attractive to members of the opposite sex (*intersexual selection*). The concept of sexual selection was either ignored or questioned for many decades, but recent investigations have demonstrated its importance.

INVESTIGATING LIFE

21.16 Sexual Selection in Male Widowbirds

The extensive tail of the territorial male African long-tailed widowbird (*Euplectes progne*) actually inhibits its ability to fly. Darwin attributed the evolution of this trait to sexual selection. Behavioral ecologist Malte Andersson tested this hypothesis.

HYPOTHESIS

Female widowbirds prefer to mate with the male that displays the longest tail; longer-tailed males thus are favored by sexual selection because they will father more offspring.

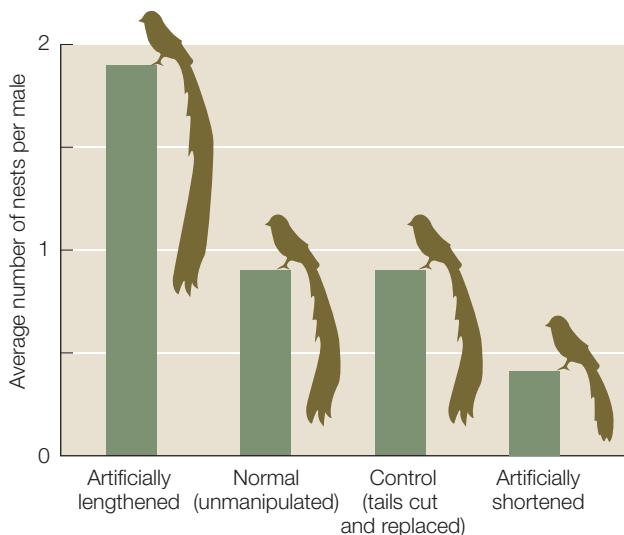
METHOD

- Capture males and artificially lengthen or shorten tails by cutting or gluing on feathers. In a control group, cut and replace tails to their normal length (to control for the effects of tail-cutting).
- Release the males to establish their territories and mate.
- Count the nests with eggs or young on each male's territory.



RESULTS

Male widowbirds with artificially shortened tails established and defended display sites successfully but fathered fewer offspring than did control or unmanipulated males. Males with artificially lengthened tails fathered the most offspring.



CONCLUSION

Sexual selection in *Euplectes progne* has favored the evolution of long tails in the male.

Go to yourBioPortal.com for original citations, discussions, and relevant links for all INVESTIGATING LIFE figures.

Darwin devoted an entire book to sexual selection because he recognized that, whereas natural selection typically favors traits that enhance the survival of their bearers or their descendants, sexual selection is primarily about success in reproduction. Of course, an animal must survive to reproduce, but if it survives and fails to reproduce, it makes no contribution to the next generation. Thus sexual selection may favor traits that enhance an individual's chances of reproduction but reduce its chances of survival. For example, females may be more likely to see or hear males with a given trait (and thus be more likely to mate with those males), even though the favored trait may also increase the chances that the male will be seen or heard by a predator. In other cases, the sexual signal may indicate a successful genotype in the male. In many species of frogs, for example, females prefer males with low-frequency calls. Males' calls vary with body size, and a low-frequency call is indicative of a large-bodied frog. Frogs exhibit indeterminate growth—that is, they continue to grow indefinitely—so a large frog is a long-lived frog, which indicates high survivorship. In this case, the sexual signal represents what is known as an *honest signal* of the male's ability to survive in the local environment.

One example of a trait that Darwin attributed to sexual selection is the remarkable tail of the male African long-tailed widowbird (*Euplectes progne*), which is longer than the bird's head and body combined. Male widowbirds normally select, and defend from other males, a territory where they perform courtship displays to attract females. To investigate whether sexual selection drove the evolution of widowbird tails, Malte Andersson, a behavioral ecologist at Gothenburg University, Sweden, clipped the tails of some captured male widowbirds and lengthened the tails of others by gluing on additional feathers. He then cut and reglued the tail feathers of still other males, which served as controls. Both short- and long-tailed males successfully defended their display territories, indicating that a long tail does not confer an advantage in male–male competition. However, males with artificially elongated tails attracted about four times more females than did males with shortened tails (Figure 21.16).

Why do female widowbirds prefer males with long tails? One possibility is that ability to grow and maintain a costly feature such as a long tail may indicate that the male bearing it is vigorous and healthy, even though the tail impairs the bird's ability to fly. If so, then females that are attracted to long tails are in-

INVESTIGATING LIFE

21.17 Do Bright Bills Signal Good Health?

Female zebra finches (*Taeniopygia guttata*) preferentially choose mates with the brightest bill color. Does this preference increase their chances of mating with the healthiest males? This experiment made use of carotenoids (antioxidant pigment molecules believed to boost immune response) to test the hypothesis.

HYPOTHESIS The brightness of a male zebra finch's red bill is correlated with the strength of the bird's immune response and a corresponding likelihood of good health.



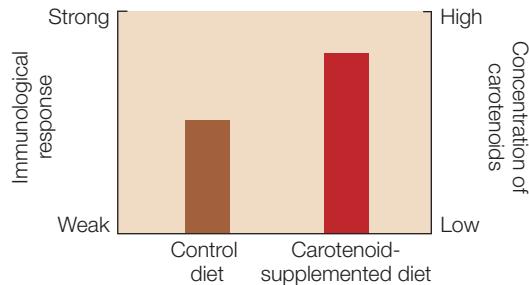
Taeniopygia guttata

METHOD

- Provide carotenoids in the drinking water of experimental, but not control, males.
- Challenge all males immunologically and measure responses.

RESULTS

Experimental males responded more strongly to the immunological challenge. They also developed brighter bills than control males.



CONCLUSION Bill color is an indication of immunological strength and general health.

FURTHER INVESTIGATION: How would you test this same hypothesis in the field? What would constitute experimental and control birds?

Go to yourBioPortal.com for original citations, discussions, and relevant links for all INVESTIGATING LIFE figures.

directly attracted to vigorous, healthy males, which likely carry beneficial genes that would lead to higher survivorship of offspring. Although the manipulated males in Andersson's investigation did not have to pay the price of growing and supporting (except briefly) artificially long tails, the hypothesis that having well-developed ornamental traits signals vigor and health has been tested experimentally with captive zebra finches.

The bright red bills of male zebra finches (*Taeniopygia guttata*) are the result of red and yellow carotenoid pigments. Zebra finches (and most other animals) cannot synthesize carotenoids and must obtain them from their food. In addition to influencing bill color, carotenoids are antioxidants and components of the immune system. Males in good health may need to allocate fewer carotenoids to immune function than males in poorer health. If so, then females can use the brightness of a male's bill to assess his health.

Tim Birkhead and his colleagues at Sheffield University manipulated blood levels of carotenoids in genetically similar male zebra finches by giving experimental males drinking water with added carotenoids; they gave control males only distilled water. All the males had access to the same food. After one month, the experimental males had higher levels of carotenoids in their blood, had much brighter bills than the control males, and were preferred by female zebra finches.

Next, the investigators challenged both groups of males immunologically by injecting phytohemagglutinin (PHA) into their wings. PHA induces a response by T lymphocytes, a type of white blood cell that functions in the immune system to recognize and deactivate foreign substances (see Chapter 42). The injection results in an accumulation of white blood cells and thus a thickening of the skin at the injection site. Experimental males with enhanced carotenoid levels developed thicker skins because they responded more strongly to PHA than control males did, indicating that higher carotenoid levels are associated with stronger immune systems (Figure 21.17).

This experiment showed that when a female chooses a male with a bright red bill, she probably gets a mate with a healthy immune system. Such males are less likely to become infected with parasites and diseases, and are better able to assist with parental care.

21.3 RECAP

Variation in genotype can lead to variation in fitness. Fitness refers to the relative reproductive contribution of a phenotype to subsequent generations. Natural and sexual selection can both change and stabilize phenotypes within populations.

- Explain why natural selection that acts on a phenotype results in changes in genotype frequencies. See p. 451
- Describe the differences between stabilizing, directional, and disruptive selection, giving examples of each. See pp. 452–453 and Figure 21.12
- Why did Darwin devote an entire book to sexual selection? See pp. 453–455

Genetic drift, stabilizing selection, and directional selection all tend to reduce genetic variation within populations. Nevertheless, as we have seen, most populations harbor considerable genetic variation. What processes produce and maintain genetic variation within populations?

21.4 How Is Genetic Variation Maintained within Populations?

Genetic variation is the raw material on which mechanisms of evolution act. In this section we will discuss several processes—neutral mutations, sexual recombination, frequency-dependent selection, and heterozygote advantage—that operate to maintain genetic variation in populations, despite the action of other forces (such as genetic drift and many types of selection) that reduce variation. We will also show how genetic variation may be maintained over geographic space.

Neutral mutations accumulate in populations

As we discuss in Section 14.6, some mutations do not affect the function of the proteins encoded by the mutated genes. An allele that does not affect the fitness of an organism—that is, an allele that is no better or worse than alternative alleles at the same locus—is called a **neutral allele**. Neutral alleles are unaffected by natural selection. Even in large populations, neutral alleles may be lost or may increase in frequency, purely by genetic drift. Neutral alleles are added to a population over time through mutation, providing the population with considerable genetic variation.

Much of the phenotypic variation we are able to observe is not neutral. However, modern techniques enable us to measure neutral variation at the molecular level and provide the means to distinguish it from adaptive variation. Section 24.2 discusses how variation in neutral molecular traits can be used to study divergence among genes, populations, and species.

Sexual recombination amplifies the number of possible genotypes

In asexually reproducing organisms, each new individual is genetically identical to its parent unless there has been a mutation. When organisms reproduce sexually, however, offspring differ from their parents because of crossing over and independent assortment of chromosomes during meiosis as well as the combination of genetic material from two different gametes, as described in Chapter 11. Sexual recombination generates an endless variety of genotypic combinations that increase the evolutionary potential of populations—a long-term advantage of sex. Although many species may reproduce asexually most of the time, few are strictly asexual over long periods of evolutionary time. Almost all have some means of achieving genetic recombination.

The evolution of the mechanisms of meiosis and sexual recombination were crucial events in the history of life. Exactly how these attributes arose is puzzling, however, because sex has at least three striking disadvantages in the short term:

- Recombination breaks up adaptive combinations of genes.
- Sex reduces the rate at which females pass genes on to their offspring.
- Dividing offspring into separate genders greatly reduces the overall reproductive rate.

To see why this last disadvantage exists, consider an asexual female that produces the same number of offspring as a sexual female. Let's assume that both females produce two offspring, but that the sexual female produces 50 percent males. In the next generation, both asexual F_1 females will produce two more offspring, but there is only one sexual F_1 female to produce offspring. Thus, the effective reproductive rate of the asexual lineage is twice that of the sexual lineage. The evolutionary problem is to identify the advantages of sex that can overcome such short-term disadvantages.

A number of hypotheses have been proposed for the existence of sex, none of which are mutually exclusive. One is that sexual recombination facilitates repair of damaged DNA, because breaks and other errors in DNA on one chromosome can be repaired by copying the intact sequence from the homologous chromosome.

Another advantage of sexual reproduction is that it permits the elimination of deleterious mutations. As Section 13.4 describes, DNA replication is not perfect. Errors are introduced in every generation, and many or most of these errors result in lower fitness. Asexual organisms have no mechanism to eliminate deleterious mutations. Hermann J. Muller noted that the accumulation of deleterious mutations in a non-recombining genome is like a genetic ratchet. The mutations accumulate—“ratchet up”—at each replication: that is, a mutation occurs and is passed on when the genome replicates, then two new mutations occur in the next replication, so three mutations are passed on, and so on. Deleterious mutations cannot be eliminated except by the death of the lineage or a rare back mutation. This accumulation of deleterious mutations in lineages that lack genetic recombination is known as *Muller's ratchet*.

In sexual species, on the other hand, genetic recombination produces some individuals with more of these deleterious mutations and some with fewer. The individuals with fewer deleterious mutations are more likely to survive. Therefore, sexual reproduction allows natural selection to eliminate particular deleterious mutations from the population over time.

Another explanation for the existence of sex is that the great variety of genetic combinations created in each generation may be advantageous. For example, genetic variation can be a defense against pathogens and parasites. Most pathogens and parasites have much shorter life cycles than their hosts and can rapidly evolve counteradaptations to host defenses. Sexual recombination might give the host's defenses a chance to keep up.

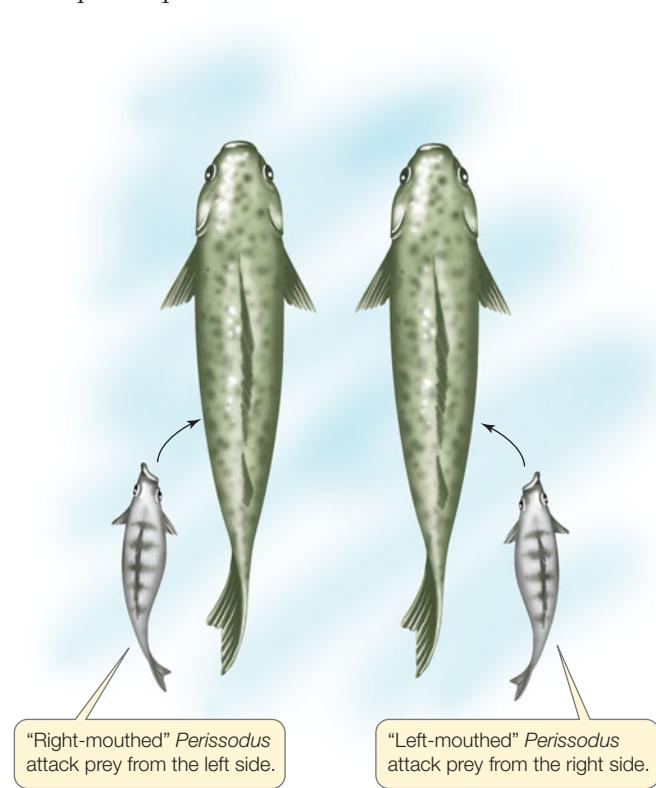
Sexual recombination does not directly influence the frequencies of alleles; rather, it generates new combinations of alleles on which natural selection can act. It expands variation in a character influenced by alleles at many loci by creating new genotypes. That is why artificial selection for bristle number in *Drosophila* (see Figure 21.5) resulted in flies that had either more or fewer bristles than the flies in the initial population had.

Frequency-dependent selection maintains genetic variation within populations

Natural selection often preserves variation as a polymorphism (two or more variants of a trait present in the same population). A polymorphism may be maintained when the fitness of a given phenotype depends on its frequency in a population, a phenomenon known as **frequency-dependent selection**.

A small fish that lives in Lake Tanganyika in East Africa provides an example of frequency-dependent selection. Because of an asymmetrical jaw joint, the mouth of this scale-eating fish, *Perissodus microlepis*, opens either to the right or to the left; the direction is genetically determined (Figure 21.18). The scale-eater approaches its prey (another fish) from behind and dashes in to bite off several scales from its flank. “Right-mouthed” individuals always attack from the victim’s left, and “left-mouthed” individuals always attack from the victim’s right. The distorted mouth enlarges the area of teeth in contact with the prey’s flank, but only if the scale-eater attacks from the appropriate side.

Prey fish are alert to approaching scale-eaters, so attacks are more likely to be successful if the prey must watch both flanks. Vigilance by prey thus favors equal numbers of right- and left-mouthed scale-eaters, because if attacks from one side were more common than the other, prey fish would pay more attention to potential attacks from that side. Over an 11-year study of this fish in Lake Tanganyika, the polymorphism was found to be stable: the two forms of *P. microlepis* remained at about equal frequencies.



21.18 A Stable Polymorphism Frequency-dependent selection maintains equal proportions of left- and right-mouthed individuals of the scale-eating fish *Perissodus microlepis*.

Heterozygote advantage maintains polymorphic loci

In many cases, different alleles for a particular gene have advantages under different environmental conditions. Most organisms, however, experience a wide diversity of environments over time. A night is dramatically different from the preceding day. A cold, cloudy day differs from a clear, hot one. Day length and temperature change seasonally. For many genes, a single allele is unlikely to perform well under all these conditions. In such situations, a heterozygous individual (with two different alleles) is likely to outperform individuals that are homozygous for either one of the alleles.

Colias butterflies of the Rocky Mountains live in environments where dawn temperatures often are too cold, and afternoon temperatures too hot, for the butterflies to fly. Populations of this butterfly are polymorphic for the enzyme phosphoglucose isomerase (PGI), which influences how well the butterfly flies at different temperatures. Butterflies with certain PGI genotypes can fly better during the cold hours of early morning; others perform better during midday heat. The optimal body temperature for flight is 35°C to 39°C, but some butterflies can fly with body temperatures as low as 29°C or as high as 40°C. During spells of unusually hot weather, heat-tolerant genotypes are favored; during spells of unusually cool weather, cold-tolerant genotypes are favored.

Heterozygous *Colias* butterflies can fly over a greater temperature range than homozygous individuals, which should give them an advantage in foraging and finding mates. A test of this prediction did find a mating advantage in heterozygous males, and further, that this advantage maintains the polymorphism in the population (Figure 21.19). Of course, the heterozygotes can never become fixed in the population, because the offspring of two heterozygotes will include both classes of homozygotes in addition to heterozygotes.

Much genetic variation in species is maintained in geographically distinct populations

Much of the genetic variation in species is preserved as differences among members living in different places (populations). Populations often vary genetically because they are subjected to different selective pressures in different environments. Environments may vary significantly over short distances. For example, in the Northern Hemisphere, temperature and soil moisture differ dramatically between north- and south-facing mountain slopes. In the Rocky Mountains of Colorado, the proportion of ponderosa pines (*Pinus ponderosa*) that are heterozygous for a particular peroxidase enzyme is particularly high on south-facing slopes, where temperatures fluctuate dramatically, often on a daily basis. This heterozygous genotype performs well over a broad range of temperatures. On north-facing slopes and at higher elevations, where temperatures are cooler and fluctuate less strikingly, a peroxidase homozygote, which has a lower optimal temperature, is much more frequent.

Plant species may also vary geographically in the chemicals they synthesize to defend themselves against herbivores. Some individuals of the white clover (*Trifolium repens*) pro-

INVESTIGATING LIFE

21.19 A Heterozygote Mating Advantage

Among butterflies of the genus *Colias*, males that are heterozygous for two alleles of the PGI enzyme can fly farther under a broader range of temperatures than males that are homozygous for either allele. Does this ability give heterozygous males a mating advantage?

HYPOTHESIS

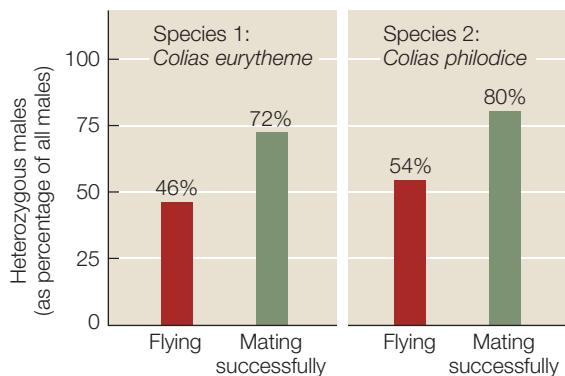
Heterozygous male *Colias* will have proportionally greater mating success than homozygous males.

METHOD

- For each of two *Colias* species, capture butterflies in the field. In the laboratory, determine their genotypes and allow them to mate.
- Determine the genotypes of the offspring, thus revealing paternity and mating success of the males.

RESULTS

For both species, the proportion of heterozygous males that mated successfully was higher than the proportion of all males seeking females ("flying").



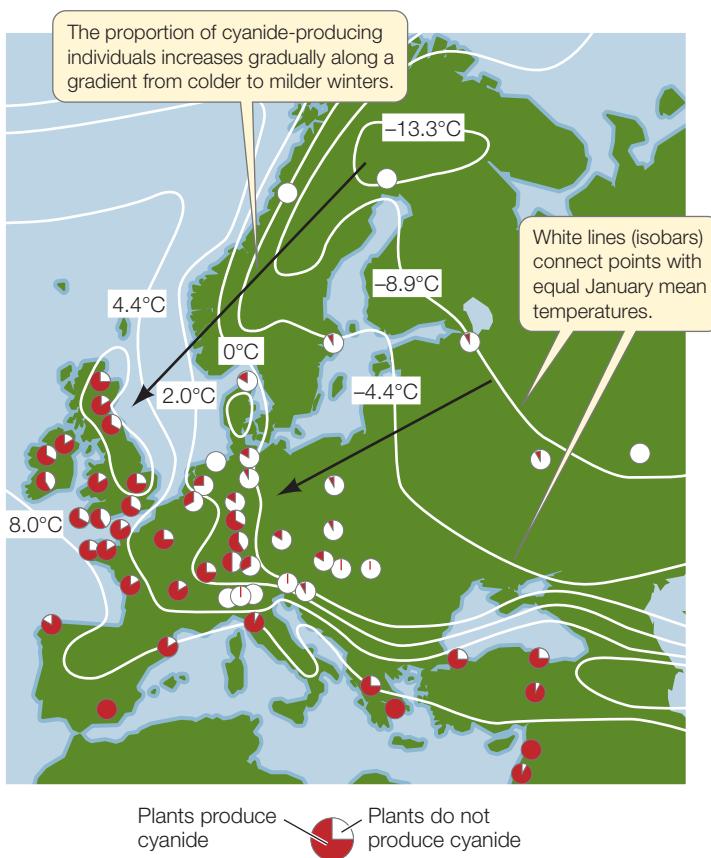
CONCLUSION

Heterozygous *Colias* males have a mating advantage over homozygous males.

Go to [yourBioPortal.com](#) for original citations, discussions, and relevant links for all INVESTIGATING LIFE figures.

duce the poisonous chemical cyanide. Poisonous individuals are less appealing to herbivores—particularly mice and slugs—than are nonpoisonous individuals. However, clover plants that produce cyanide are more likely to be killed by frost, because freezing damages cell membranes and releases cyanide into the plant's own tissues.

In European populations of *Trifolium repens*, the frequency of cyanide-producing individuals increases gradually from north to south and from east to west (Figure 21.20). This gradual change in phenotype across a geographic gradient is known as **clinal variation**. In this cline, poisonous plants make up a large proportion of clover populations only in areas where winters are mild. Cyanide-producing individuals are rare where winters are cold, even though herbivores graze clovers heavily in those areas.



21.20 Geographic Variation in a Defensive Chemical

The frequency of cyanide-producing individuals in European populations of white clover (*Trifolium repens*) depends on winter temperatures.

21.4 RECAP

Neutral mutations, sexual recombination, frequency-dependent selection, and heterozygote advantage all act to maintain considerable genetic variation in most populations. Variation within species is also maintained among geographically distinct, genetically variable populations.

- Do you understand why sexual reproduction is so prevalent in nature, despite its having at least three short-term evolutionary disadvantages? [See p. 456](#)
- How does frequency-dependent selection act to maintain genetic variation in a population? [See p. 457](#)

The mechanisms of evolution have produced a remarkable variety of organisms, some of which are adapted to most environments on Earth. This natural variation, and the success of breeders attempting to produce desired traits in domesticated plants and animals, suggests that evolution can produce a wide variety of adaptive traits. But are there limits to the adaptations evolution can produce?

21.5 What Are the Constraints on Evolution?

We would be mistaken to assume that evolutionary mechanisms can produce any trait we might imagine. Evolution is constrained in many ways. Lack of appropriate genetic variation, for example, prevents the development of many potentially favorable traits. If the allele for a given trait does not exist in a population, that trait cannot evolve even if it would be highly favored by natural selection. Most possible combinations of genes and genotypes have never existed in any population, and so have never been tested under natural selection.

Constraints are imposed on organisms by the dictates of physics and chemistry. The size of cells, for example, is constrained by the stringencies of surface area-to-volume ratios (see Section 2.1). The ways in which proteins can fold are limited by the bonding capacities of their constituent molecules (see Section 3.2). And the energy transfers that fuel life must operate within the laws of thermodynamics (see Section 8.1). Keep in mind that evolution works within the boundaries of these universal constraints, as well as the constraints described here.

Developmental processes constrain evolution

As Section 20.5 explained, developmental constraints on evolution are paramount because *all evolutionary innovations are modifications of previously existing structures*. Human engineers seeking to power an airplane can start “from scratch” to design a completely new type of engine (powered by jet propulsion), to replace the previous type (powered by propellers). Evolutionary changes, however, cannot happen in this way. Current phenotypes of organisms are constrained by historical conditions and past selective pressures.

A striking example of such developmental constraints is provided by the evolution of fish that spend most of their time on the sea bottom. One lineage, the bottom-dwelling skates and rays, share a common ancestor with sharks, whose bodies were already somewhat ventrally flattened and whose skeletal frame is made of flexible cartilage. Skates and rays evolved a body type that further flattened their bellies, allowing them to swim along the ocean floor (Figure 21.21A).

By contrast, plaice, sole, and flounder are bottom-dwelling descendants of deep-bellied, laterally flattened ancestors with bony skeletons. The only way these fishes can lie flat is to flop over on their sides. Their ability to swim is thus curtailed, but their bodies can lie still and are well camouflaged. During development, one eye of these flatfishes moves so that both eyes are positioned on the same side of the body (Figure 21.21B). Such shifts in eye position have evolved several times, and shifts have happened in both directions (that is, both left- and right-eyed flatfishes have evolved independently). Small shifts in the

21.21 Two Solutions to a Single Problem (A) This stingray, whose ancestors were dorsoventrally flattened, lies on its belly. Stingrays' bodies are symmetrical around the dorsal backbone. (B) This flounder, whose ancestors were laterally flattened, lies on its side. (The backbone of this individual is at the right.) Flounders' eyes migrate during development so that both are on the same side of the body.

position of one eye probably helped ancestral flatfishes see better, resulting in the flat body forms found today. This path to producing a flattened body may not be optimal, but the fishes' developmental capabilities constrain the pathways that evolution can take.

Trade-offs constrain evolution

Adaptations frequently impose both fitness costs and benefits. For an adaptation to evolve, the fitness benefits it confers must exceed the fitness costs it imposes—in other words, the **trade-off** must be worthwhile. For example, there are metabolic costs associated with developing and maintaining certain conspicuous features (such as antlers or horns) that males use to compete with other males for access to females. The fact that these features are common in many species suggests that the benefits derived from possessing them must outweigh the costs.

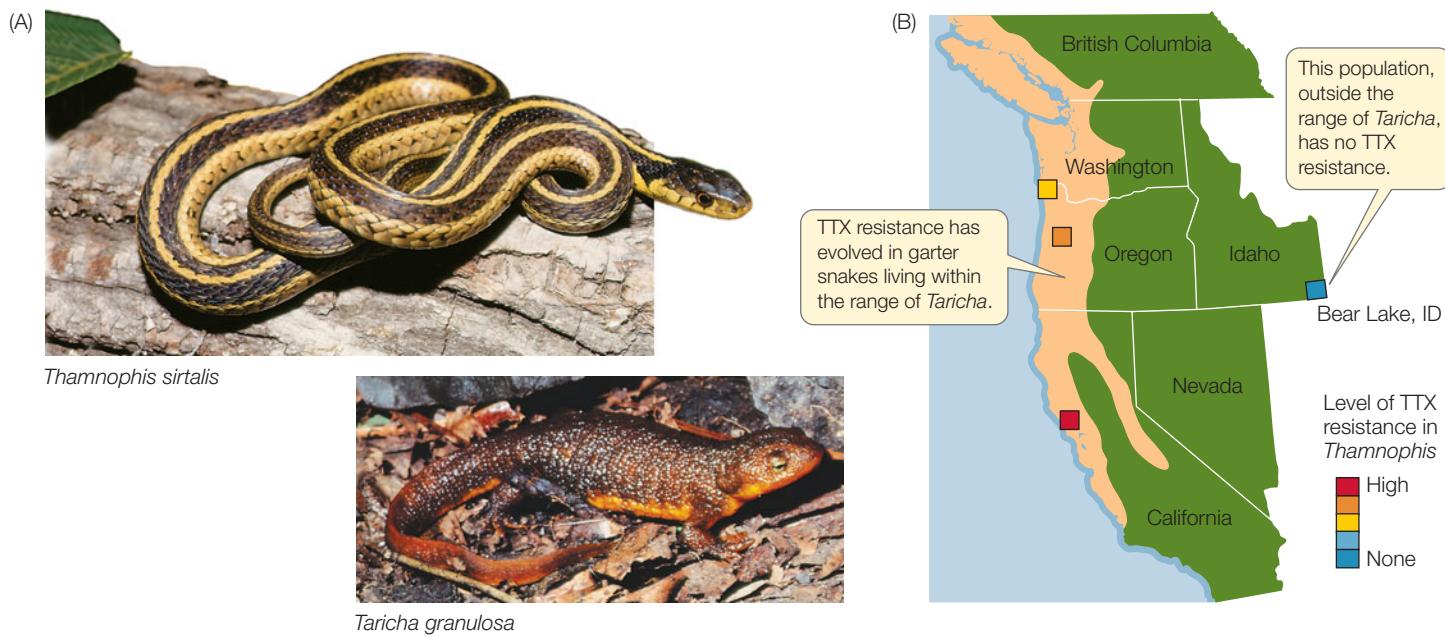
As a result of trade-offs, many traits that are adaptive in one context may be maladaptive in another. Consider the rough-skinned newt, *Taricha granulosa*, and one of its predators, the common garter snake, *Thamnophis sirtalis* (Figure 21.22A). The newt sequesters in its skin a potent neurotoxin called tetrodotoxin (TTX). TTX paralyzes nerves and muscles by blocking



(A) *Taeniura lymma*



(B) *Bothus lunatus*



21.22 Resistance to a Toxin Comes at a Cost (A) Garter snakes (above) prey on newts (below). Rough-skinned newts counter with the ability to sequester a neurotoxin, TTX, in their skin. In turn, TTX-resistant sodium channels have evolved in some snake populations, allowing the snakes to eat toxic prey but resulting in slower movement by the snakes. (B) High resistance to TTX in garter snakes is only found in regions where snake and newt populations overlap (tan area).

yourBioPortal.com
GO TO Animated Tutorial 21.3 •
Assessing the Costs of Adaptation

sodium channels (see Section 6.3). Most vertebrates—including many garter snakes—will die if they eat a rough-skinned newt. But some snakes can eat rough-skinned newts and survive. In some populations of garter snakes, TTX-resistant sodium channels have evolved in the nerves and muscles (see Chapter 24 for another example of the evolution of sodium channels). However, the snakes pay a price for this attribute. For several hours after eating a newt, TTX-resistant snakes can move only slowly, and they never move as fast as nonresistant snakes. Thus resistant snakes are more vulnerable to their own predators than are TTX-sensitive snakes that simply don't encounter poisonous newts. Therefore, there is selection against TTX-resistant sodium channels in populations of garter snakes that occur outside the range of rough-skinned newts, but selection for TTX-resistance in many areas where newts are present (Figure 21.22B).

Short-term and long-term evolutionary outcomes sometimes differ

The short-term changes in allele frequencies within populations that we have emphasized in this chapter are an important focus of study for evolutionary biologists. These changes can be

observed directly, they can be manipulated experimentally, and they demonstrate the actual processes by which evolution occurs. By themselves, however, they do not enable us to predict long-term evolutionary changes.

Long-term patterns of evolutionary change can be strongly influenced by events that occur so infrequently (a meteorite impact, for example) or so slowly (continental drift) that they are unlikely to be observed during short-term studies. The ways in which evolutionary processes act may change over time with changing environmental conditions. Even among the descendants of a single ancestral species, different lineages may evolve in different directions. Therefore, additional types of evidence, demonstrating the effects of rare and unusual events on trends in the fossil record, must be gathered if we wish to understand the course of evolution over billions of years.

21.5 RECAP

Developmental processes constrain evolution because all evolutionary innovations are modifications of previously existing structures. An adaptation can evolve only if the fitness benefits it confers exceed the fitness costs it imposes.

- Describe an example of an evolutionary trade-off in which the advantages of an adaptation outweigh its costs in the long run. See pp. 459–460
- Do you see why the presence of a great deal of genetic variation within a population could increase the chances that some members of the population would survive an unprecedented environmental change? Do you also understand why there is no guarantee that this would be the case?

CHAPTER SUMMARY

21.1 What Facts Form the Base of Our Understanding of Evolution?

- Charles Darwin attributed changes in species over time to the possession of advantageous traits by some individuals. He understood that it is not individuals that evolve but **populations**. A population evolves when individuals with favorable **heritable traits** survive and reproduce at higher rates than other members of the population.
- Adaptation** refers both to characteristics of organisms and the way those characteristics are acquired via **natural selection**.
ANIMATED TUTORIAL 21.1
- The sum of all copies of all alleles at all loci found in a population constitutes its **gene pool** and represents the genetic variation that results in different phenotypic traits on which natural selection can act. **Review Figure 21.3**
- Artificial selection** and laboratory experiments demonstrate the existence of considerable genetic variation in most populations. **Review Figure 21.5**
- Allele frequencies measure the amount of genetic variation in a population; genotype frequencies show how a population's genetic variation is distributed among its members. Together, allele and genotype frequencies describe a population's **genetic structure**. **Review Figure 21.6**
- Hardy–Weinberg equilibrium** predicts the allele frequencies in populations in the absence of evolution. Deviation from these frequencies indicates the work of evolutionary mechanisms. **Review Figure 21.7, ANIMATED TUTORIAL 21.2**

21.2 What Are the Mechanisms of Evolutionary Change?

- Mutation** provides new genetic variants; favored variants increase in populations through natural selection.
- Migration or mating of individuals between populations results in **gene flow**.
- In small populations, **genetic drift**—the random loss of individuals and the alleles they possess—may produce large changes in allele frequencies from one generation to the next and greatly reduce genetic variation. **Review Figure 21.8**
- Population bottlenecks** occur when only a few individuals survive a random event, resulting in a drastic shift in allele frequencies within the population and the loss of variation. Similarly, a population established by a small number of individuals colonizing a new region may lose variation via a **founder effect**.
- Nonrandom mating may result in genotype frequencies that deviate from Hardy–Weinberg equilibrium.

21.3 How Does Natural Selection Result in Evolution?

- Fitness** is the reproductive contribution of a phenotype to subsequent generations relative to the contributions of other phenotypes.
- Changes in numbers of offspring are responsible for changes in the absolute size of a population, but only changes in the relative success of different phenotypes within a population lead to changes in allele frequencies.
- Natural selection can act on variable traits in several different ways, resulting in **stabilizing**, **directional**, or **disruptive selection**. **Review Figure 21.12**
- Sexual selection** primarily affects success in reproduction, rather than success in survival. **Review Figures 21.16 and 21.17**

21.4 How Is Genetic Variation Maintained within Populations?

- Neutral mutations, sexual recombination, frequency-dependent selection, and heterozygote advantage can all maintain genetic variation within populations.
- Neutral alleles** do not affect the fitness of an organism, are not affected by natural selection, and may accumulate or be lost by genetic drift.
- Despite short-term disadvantages, sexual reproduction generates countless genotypic combinations that increase the evolutionary potential and survivorship of populations.
- A polymorphism may be maintained by **frequency-dependent selection** when the fitness of a genotype depends on its frequency in a population.
- Genetic variation within species may be maintained by the existence of genetically distinct populations over geographic space. A gradual change in phenotype across a geographic gradient is known as **clinal variation**. **Review Figure 21.20**

21.5 What Are the Constraints on Evolution?

- Developmental processes constrain evolution because all evolutionary innovations are modifications of previously existing structures.
- Most adaptations impose costs. An adaptation can evolve only if the benefits it confers exceed the costs it imposes, a situation that leads to **trade-offs**. **Review Figure 21.22, ANIMATED TUTORIAL 21.3**

SELF-QUIZ

- Long-horned cattle have greater difficulty moving through heavily forested areas compared with cattle that have short or no horns, but long-horned cattle are better able to defend their young against predators. This contrast is an example of
 - an adaptation.
 - genetic drift.
 - natural selection.
 - a trade-off.
 - none of the above
- Which of the following is true?
 - Darwin and Wallace were both influenced by Malthus.
 - Wallace proposed a theory of evolution by natural selection that was similar to Darwin's.
 - Malthus claimed that because human population growth would outstrip any increases in food production, famine was a likely result.
 - Darwin realized that all populations had the capacity to rapidly increase in numbers.
 - All of the above

3. The phenotype of an organism is
 - a. the type specimen of its species in a museum.
 - b. its genetic constitution, which governs its traits.
 - c. the chronological expression of its genes.
 - d. the physical expression of its genotype.
 - e. its adult form.
4. The appropriate unit for defining and measuring genetic variation is the
 - a. cell.
 - b. individual.
 - c. population.
 - d. community.
 - e. ecosystem.
5. Which statement about allele frequencies is *not* true?
 - a. The sum of all allele frequencies at a locus is always 1.
 - b. If there are two alleles at a locus and we know the frequency of one of them, we can obtain the frequency of the other by subtraction.
 - c. If an allele is missing from a population, its frequency in that population is 0.
 - d. If two populations have the same allele frequencies at a locus, they must have the same proportion of homozygotes at that locus.
 - e. If there is only one allele at a locus, its frequency is 1.
6. Which of the following is *not* required for a population at Hardy–Weinberg equilibrium?
 - a. There is no migration between populations.
 - b. Natural selection is not acting on the alleles in the population.
 - c. Mating is random.
 - d. Multiple alleles must be present at every locus.
 - e. All of the above.
7. The fitness of a genotype is a function of the
 - a. average rates of survival and reproduction of individuals with that genotype.
 - b. individuals that have the highest rates of both survival and reproduction.
8. Laboratory selection experiments with fruit flies have demonstrated that
 - a. bristle number is not genetically controlled.
 - b. bristle number is not genetically controlled, but changes in bristle number are caused by the environment in which the fly is raised.
 - c. bristle number is genetically controlled, but there is little variation on which natural selection can act.
 - d. bristle number is genetically controlled, but selection cannot result in flies having more bristles than any individual in the original population had.
 - e. bristle number is genetically controlled, and selection can result in flies having more, or fewer, bristles than any individual in the original population had.
9. Disruptive selection maintains a bimodal distribution of bill size in the West African seedcracker because
 - a. bills of intermediate shapes are difficult to form.
 - b. the birds' two major food sources differ markedly in size and hardness.
 - c. males use their large bills in displays.
 - d. migrants introduce different bill sizes into the population each year.
 - e. older birds need larger bills than younger birds.
10. Which of the following is *not* a reason why trade-offs constrain evolution?
 - a. Most adaptations impose both fitness costs and benefits.
 - b. Structures such as horns and antlers are metabolically costly to produce, but result in more reproduction by the males that possess them.
 - c. Changes in allele frequencies may be influenced by chance events.
 - d. Ability to consume toxic prey may reduce mobility.
 - e. Adaptations can evolve only if the fitness benefits they confer exceed the costs they impose.

FOR DISCUSSION

1. In what ways does artificial selection by humans differ from natural selection? Was Darwin wise to base so much of his argument for natural selection on the results of artificial selection?
2. In nature, mating among individuals in a population is never truly random; immigration and emigration are common, and natural selection is continuous. Why, then, is Hardy–Weinberg equilibrium, which is based on assumptions known generally to be false, so useful in our study of evolution? Can you think of other models in science that are based on false assumptions? How are such models used?
3. As far as we know, natural selection cannot adapt organisms to future events. Yet many organisms appear to respond to natural events before they happen. For example, many mammals go into hibernation while it is still quite warm. Similarly, many birds leave the temperate zone for their southern wintering grounds long before winter has arrived. How can such “anticipatory” behaviors evolve?
4. Populations of most of the thousands of species that have been introduced to areas where they were previously not found, including those that have become pests, began with a few individuals. Founding populations therefore begin with much less genetic variation than their parental populations have. If genetic variation is generally advantageous, why have so many of these species been successful in their new environments?
5. Why is it important that the ways in which males advertise their health and vigor to females reliably indicate their status?
6. As more humans live longer, many people face degenerative conditions such as Alzheimer’s disease that (in most cases) are linked to advancing age. Assuming that some individuals may be genetically predisposed to successfully combat these conditions, is it likely that natural selection alone would act to favor such a predisposition in human populations? Why or why not?

ADDITIONAL INVESTIGATION

During the past 50 years, more than 200 species of insects that attack crop plants have become highly resistant to DDT and other pesticides. Using your recently acquired knowledge of evolutionary processes, explain the rapid and widespread evolution

of resistance. What proposals concerning pesticide use would you make in order to slow down the rate of evolution of resistance? Explain why you think your proposals could work and how you might test them.

WORKING WITH DATA (GO TO yourBioPortal.com)

Testing for Significant Differences In this hands-on exercise based on Figure 21.16, you will use a simple method for randomizing Malte Andersson's data to test for significant differences among the various experimental groups. You will also explore how sample size affects the power to make significant conclusions in experiments.

Female Mating Preference in Zebra Finches In this exercise based on Figure 21.17, you will evaluate the data Jonathan Blount and his colleagues used to demonstrate female preference for males with bright bills among zebra finches. You will also consider the limitations to the experiment and explore alternative study designs.

Determining the Paternity of Butterfly Larvae In working with a sample of the data collected by Ward et al. for the experiment described in Figure 21.19, you will consider how many larvae from a clutch of butterfly eggs must be examined in order to determine (with a high level of confidence) whether the clutch was fathered by a heterozygous or a homozygous male. You will also consider alternative hypotheses to the authors' conclusions and suggest how these alternative ideas could be tested.

Phylogenetic trees in the courtroom

Transmitting HIV, while irresponsible, is not usually prosecuted as a crime. But in one true-crime case, a woman we'll call "April" went to the police immediately upon learning she was HIV-positive. April believed she was the victim of an attempted murder by "Victor," a physician and her former boyfriend, who had repeatedly threatened violence when she tried to break up with him. April contended that Victor, under the pretense of administering vitamin therapy, had injected her with blood from one of his HIV-infected patients.

Police investigators discovered that Victor had drawn blood from one of his HIV-positive patients just before giving April the injection. The blood draw had no clinical

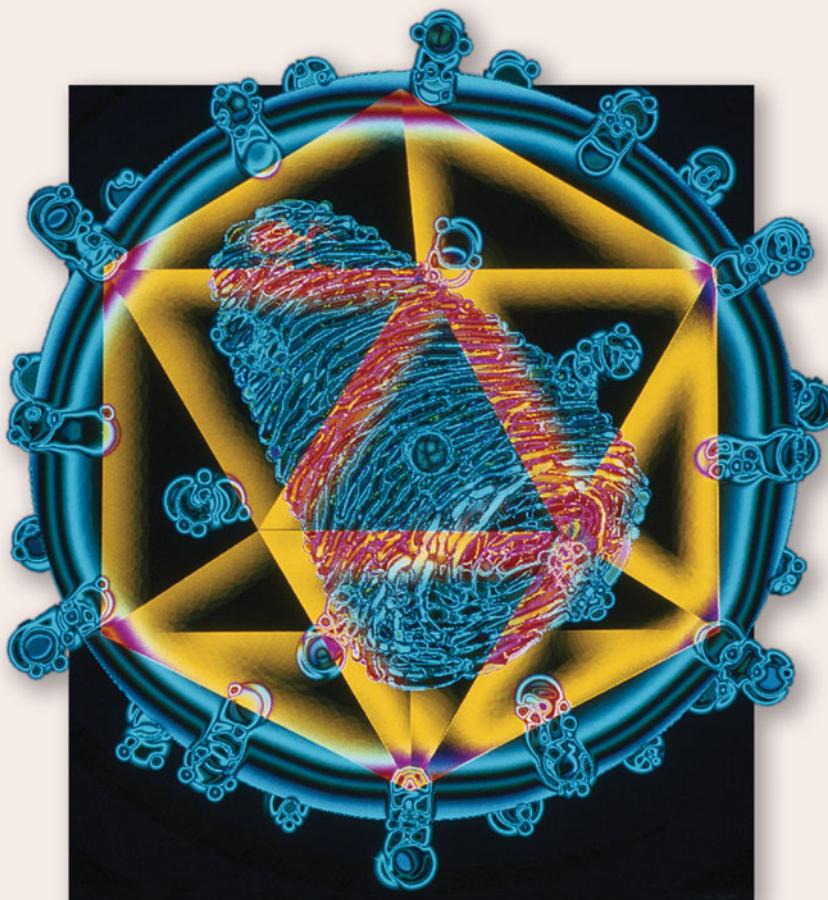
purpose, and Victor had tried to hide the records of it. The police were convinced he might indeed have committed the alleged crime.

The district attorney, however, had to show that April's HIV infection had come from Victor's patient, and from no other source. To reconstruct the history of the infection, the district attorney turned to *phylogenetic analysis*—the study of the evolutionary relationships among a group of organisms.

The district attorney's task was complicated by the nature of HIV. HIV is a retrovirus, in which poor repair of replication errors leads to a high rate of evolution. Once a person is infected with HIV, the virus not only replicates quickly but evolves quickly, so that the infected individual is soon host to a genetically diverse population of viruses. Thus when one person transmits HIV to another, typically

very few viral particles (often only one) initiate the infectious event. But the person who is the source of the infection may be host to a large, genetically diverse population of viruses—not just the variant he or she transmits to the recipient.

Enter molecular phylogeny. Samples of HIV from an infected individual can be sequenced to trace their evolutionary lineages back to the originally transmitted virus. The virus that is passed to the recipient will be very closely related to some of the viruses in the source individual and more distantly related to others. A reconstruction of the evolutionary history of the viruses in both individuals is needed to reveal not only whether the two individuals' viruses are closely related, but also who infected whom.



Human Immunodeficiency Virus A computer image of the human immunodeficiency virus (HIV), the cause of acquired immunodeficiency syndrome, or AIDS. To combat AIDS, it is also essential to understand the phylogeny of HIV.



A Source of the Virus AIDS is a zoonotic disease, meaning that the virus was transferred to humans from another animal. Phylogenetic analyses of immunodeficiency viruses show that humans acquired HIV-1 from chimpanzees (see Figure 22.9). Other forms of the virus have been passed to humans by different simians.

To prove attempted murder, the district attorney needed to demonstrate that April's HIV was more closely related to that of Victor's patient than to other HIV variants in her community. Samples of HIV were isolated from the blood of the patient, from April, and from other HIV-positive individuals in the community. Phylogenetic analysis revealed that April's HIV was indeed closely related to a subset of the patient's HIV, and more distantly related to the other HIV sources in the community. Given this fact, along with other evidence in the case, Victor was convicted of attempted murder.

IN THIS CHAPTER we will examine the field of systematics, the scientific study of the diversity of life. We will see how phylogenetic methods are used to reconstruct evolutionary history and to study diversity across genes, populations, species, and larger groups of organisms. We will see how systematists reconstruct the past and use phylogenies to make predictions in biology. We will end the chapter with a look at taxonomy, the theory and practice of classifying organisms.

CHAPTER OUTLINE

- 22.1** What Is Phylogeny?
- 22.2** How Are Phylogenetic Trees Constructed?
- 22.3** How Do Biologists Use Phylogenetic Trees?
- 22.4** How Does Phylogeny Relate to Classification?

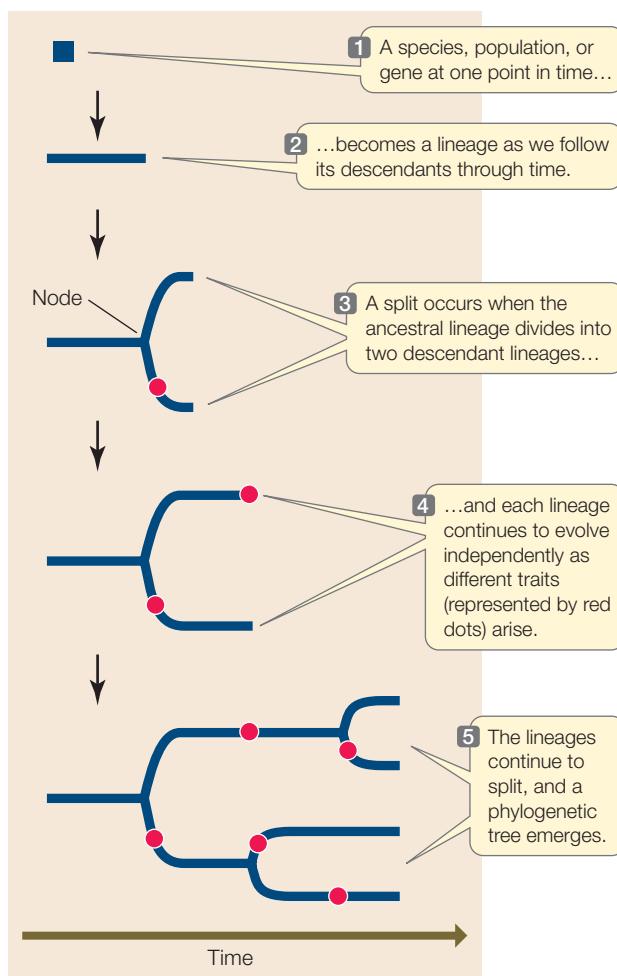
22.1 What Is Phylogeny?

Phylogeny is the evolutionary history of relationships among organisms or their genes. A **phylogenetic tree** is a diagram that portrays a reconstruction of that history. Phylogenetic trees are commonly used to depict the evolutionary history of species, populations, and genes. Each split (or *node*) in a phylogenetic tree represents a point at which lineages diverged in the past. In the case of species, these splits represent past speciation events, when one lineage divided into two. Thus a phylogenetic tree can be used to trace the evolutionary relationships from the ancient common ancestor of a group of species, through the various speciation events (when lineages split), up to the present populations of the organisms (Figure 22.1). Over the past several decades, phylogenetic trees have become important tools for studying and describing evolutionary patterns, and for applying evolutionary theory throughout biology. You will need to understand phylogenetic trees to comprehend many articles and books about biology, including this one.

A phylogenetic tree may portray the evolutionary history of all life forms; of a major evolutionary group (such as the insects); of a small group of closely related species; or in some cases, even the history of individuals, populations, or genes within a species. The common ancestor of all the organisms in the tree forms the *root* of the tree. The phylogenetic trees in this book depict time flowing from left (earliest) to right (most recent) (Figure 22.2A). It is also common practice to draw trees with the earliest times at the bottom.

The timing of splitting events in lineages is shown by the position of nodes on a time axis, sometimes called a *divergence axis*. These splits represent events where one lineage diverged into two, such as a speciation event (for a tree of species), a gene duplication event (for a tree of genes), or a transmission event (for a tree of viral lineages transmitted through a host population). The divergence axis may have an explicit scale or simply show the relative timing of splitting events. In this book, the order of nodes along the horizontal (time) axis have meaning, but the vertical distance between the branches does not. Vertical distances are adjusted for legibility and clarity of presentation; they do not correlate with the degree of similarity or difference between groups. Note too that lineages can be rotated around nodes in the tree, so the vertical order of lineages is also largely arbitrary (Figure 22.2B). The important information in the tree is the branching order along the horizontal axis, as this indicates when the various lineages last shared a common ancestor.

Any group of species that we designate or name is called a **taxon** (plural *taxa*). Some examples of familiar taxa include humans, primates, mammals, and vertebrates (note that in this



22.1 A Phylogenetic Tree Evolutionary relationships among lineages, as well as the evolution of new traits, can be represented in a treelike diagram.

series, each taxon in the list is also a member of the next, more inclusive taxon). Any taxon that consists of all the evolutionary descendants of a common ancestor is called a **clade**. Clades can be identified by picking any point on a phylogenetic tree and then tracing all the descendant lineages to the tips of the terminal branches (Figure 22.3). Two species that are each other's closest relatives are called **sister species**; similarly, any two clades that are each other's closest relatives are called **sister clades**.

Before the 1980s, phylogenetic trees tended to be seen only in the literature on evolutionary biology, especially in the area of **systematics**: the study and classification of biodiversity. But almost every journal in the life sciences published during the last few years con-

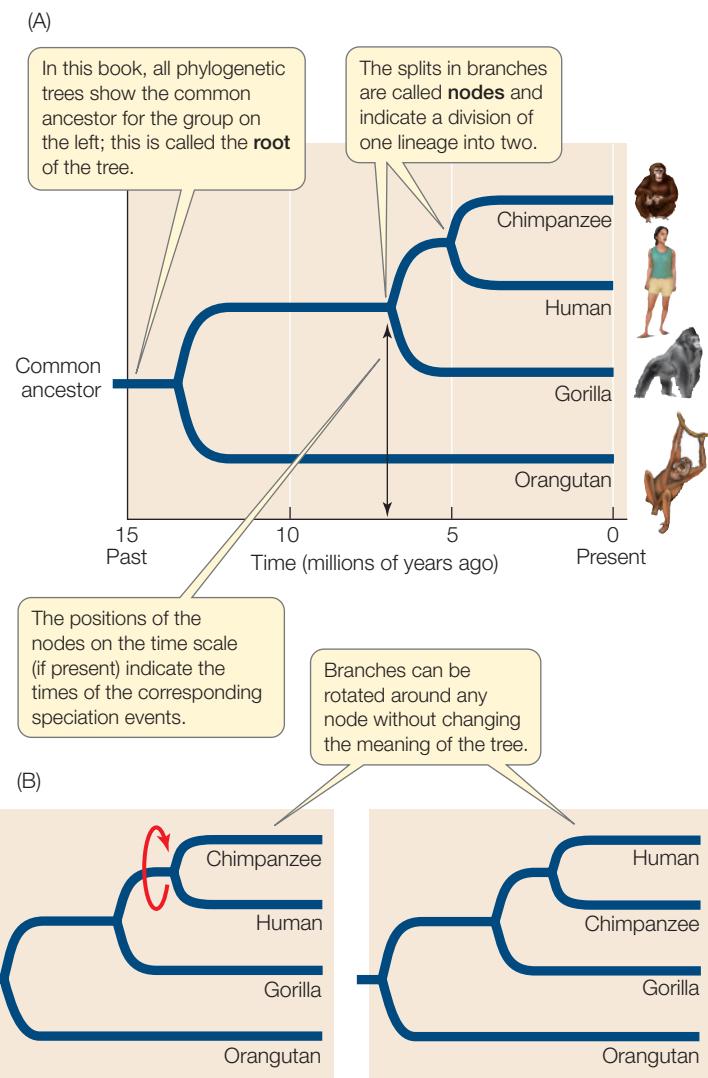
22.2 How to Read a Phylogenetic Tree (A) A phylogenetic tree displays the evolutionary relationships among organisms. Such trees can be produced with time scales, as shown here, or with no indication of time. If no time scale is shown, then the branch lengths show relative rather than absolute times of divergence. (B) Lineages can be rotated around a given node, so the vertical order of taxa is largely arbitrary.

tains phylogenetic trees. Trees are widely used in molecular biology, biomedicine, physiology, behavior, ecology, and virtually all other fields of biology. Why have phylogenetic studies become so important?

All of life is connected through evolutionary history

In biology, we study life at all levels of organization—from genes, cells, organisms, populations, and species to the major divisions of life. In most cases, however, no individual gene or organism (or other unit of study) is exactly like any other gene or organism that we investigate.

Consider the individuals in your biology class. We recognize each person as an individual human, but we know that no two are exactly alike. If we knew everyone's family tree in detail, the genetic similarity of any pair of students would be more predictable. We would find that more closely related students have many more traits in common (from the color of their hair to their susceptibility or resistance to diseases). Likewise, biologists use phylogenies to make comparisons and predictions about shared traits across genes, populations, and species.



22.3 Clades Represent All the Descendants of a Common Ancestor

Ancestor All clades are subsets of larger clades, with all of life as the most inclusive taxon. In this example, the groups called mammals, amniotes, tetrapods, and vertebrates represent successively larger clades. Only a few species within each clade are represented on the tree.

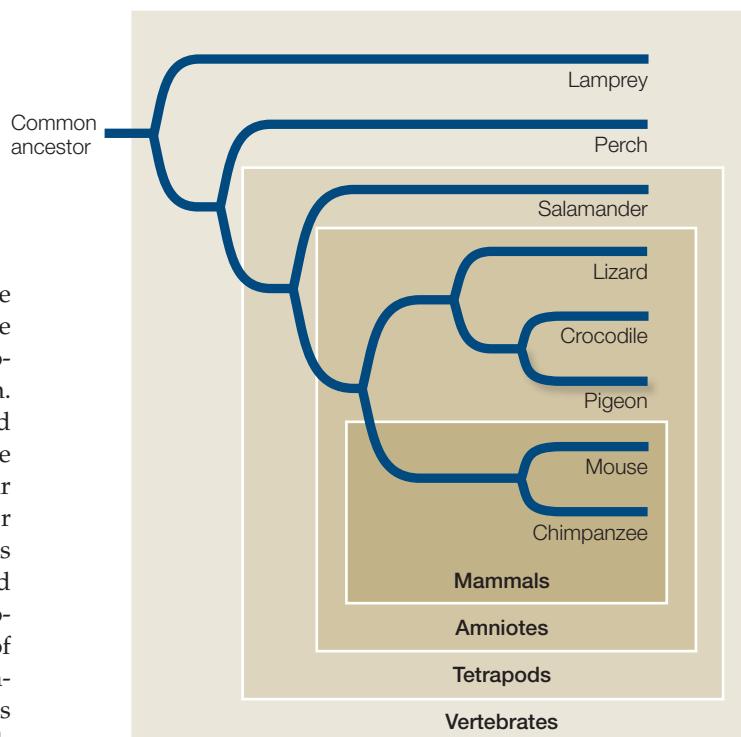
One of the great unifying concepts in biology is that all life is connected through its evolutionary history. The complete evolutionary history of life is known as the **tree of life**. Biologists estimate that there are tens of millions of species on Earth. Only about 1.8 million have been formally described and named. New species are being discovered and named all the time, and phylogenetic analyses reviewed and revised, but our knowledge of the tree of life is far from complete, even for known species. Yet knowledge of evolutionary relationships is essential for making comparisons in biology, so biologists build phylogenies for groups of interest as the need arises. The evolutionary relationships among species, as shown in the tree of life, form the basis for biological classification. This evolutionary framework allows biologists to make many predictions about the behavior, ecology, physiology, genetics, and morphology of species that have not yet been studied in detail.

Comparisons among species require an evolutionary perspective

When biologists make comparisons among species, they observe traits that differ within the group of interest and try to ascertain when these traits evolved. In many cases, investigators are interested in how the evolution of a trait depends on environmental conditions or selective pressures. For instance, scientists have used phylogenetic analyses to discover changes in the genome of HIV that confer resistance to particular drug treatments. The association of a particular genetic change in HIV with a particular treatment provides a hypothesis about the evolution of resistance that can be tested experimentally.

Any features shared by two or more species that have been inherited from a common ancestor are said to be **homologous**. Homologous features may be any heritable traits, including DNA sequences, protein structures, anatomical structures, and even some behavior patterns. Traits that are shared across a group of interest are likely to have been inherited from a common ancestor. For example, all living vertebrates have a vertebral column, and all known fossil vertebrates had a vertebral column. Therefore, the vertebral column is judged to be homologous in all vertebrates.

In tracing the evolution of a trait, biologists distinguish between *ancestral* and *derived* traits. A trait that was already present in the ancestor of a group is known as an **ancestral trait** for that group. A trait found in a descendent that differs from its ancestral form is called a **derived trait**. Derived traits that are shared among a group of organisms, and are viewed as evidence of the common ancestry of the group, are called **synapomorphies** (*syn*, “shared”; *apo*, “derived”; *morph*, “form,” referring to the “form” of a trait). Thus the vertebral column is



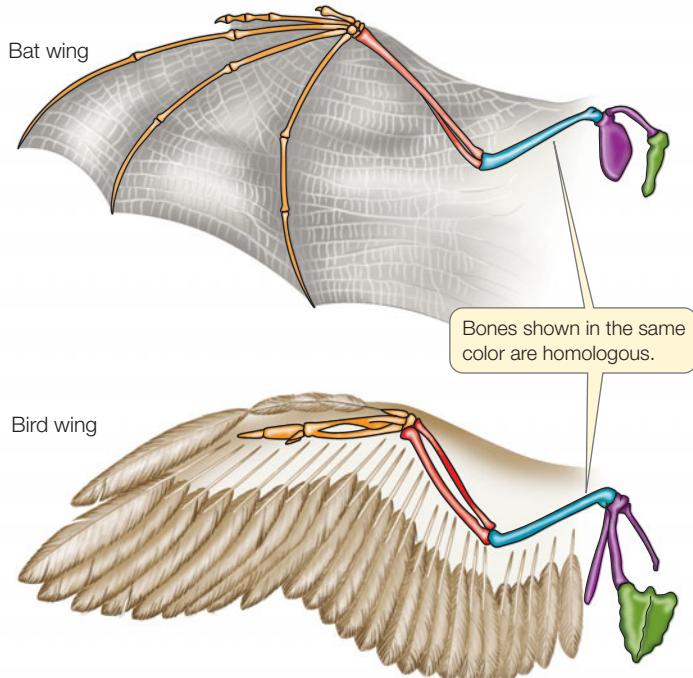
considered a synapomorphy—a shared, derived trait—of the vertebrates.

Not all similar traits are evidence of relatedness, however. Similar traits in unrelated groups of organisms can develop for either of the following reasons:

- Independently evolved traits subjected to similar selection pressures may become superficially similar, a phenomenon called **convergent evolution**. For example, although the wing bones of bats and birds are homologous, having been inherited from a common ancestor, the wings of bats and the wings of birds are not homologous because they evolved independently from the forelimbs of different non-flying ancestors (Figure 22.4).
- A character may revert from a derived state back to an ancestral state in an event called an **evolutionary reversal**. For example, most frogs lack teeth in the lower jaw, but the ancestor of frogs did have such teeth. Teeth have been regained in the lower jaw of one South American species, and thus represent an evolutionary reversal in that species.

Similar traits generated by convergent evolution and evolutionary reversals are called **homoplastic traits** or **homoplasies**.

A particular trait may be ancestral or derived, depending on our point of reference in a phylogeny. For example, all birds have feathers, which are highly modified scales. We infer from this that feathers were present in the common ancestor of modern birds. Therefore, we consider the presence of feathers to be an *ancestral* trait for any particular group of modern birds, such as the songbirds. However, feathers are not present in any other living animals. If we were reconstructing a phylogeny of all



22.4 The Bones Are Homologous, the Wings Are Not The supporting bone structures of both bat wings and bird wings are derived from a common four-limbed ancestor and are thus homologous. However, the wings themselves—an adaptation for flight—evolved independently in the two groups.

living vertebrates, the presence of feathers would be a *derived* trait that is found only among birds (and thus a synapomorphy of the birds).

22.1 RECAP

A phylogenetic tree is a description of evolutionary relationships—how a group of genes, populations, or species have evolved from a common ancestor. All living organisms share a common ancestor and are related through the phylogenetic tree of life.

- Do you understand the different elements of a phylogenetic tree? **See pp. 465–466 and Figure 22.2**
- Explain the difference between an ancestral and a derived trait. **See p. 467**
- Do you see how similar traits might arise independently in species that are only distantly related? **See p. 467 and Figure 22.4**

Phylogenetic analyses have become increasingly important to many types of biological research in recent years, and they are the basis for the comparative nature of biology. For the most part, however, evolutionary history cannot be observed directly. How, then, do biologists reconstruct the past? One way is by using phylogenetic analyses to construct a tree.

22.2 How Are Phylogenetic Trees Constructed?

To illustrate how a phylogenetic tree is constructed, let's consider the eight vertebrate animals listed in **Table 22.1**: lamprey, perch, salamander, lizard, crocodile, pigeon, mouse, and chimpanzee. We will assume initially that a given derived trait evolved only once during the evolution of these animals (that is, there has been no convergent evolution), and that no derived traits were lost from any of the descendant groups (there has been no evolutionary reversal). For simplicity, we have selected traits that are either present (+) or absent (-).

In a phylogenetic study, the group of organisms of primary interest is called the **ingroup**. As a point of reference, an ingroup is compared with an **outgroup**, a closely related species or group known to be phylogenetically outside the group of interest. If the outgroup is known to have diverged before the ingroup, the outgroup can be used to determine which traits of the ingroup are derived (evolved within the ingroup) and which are ancestral (evolved before the origin of the ingroup). As we will see in Chapter 33, a group of jawless fishes called the lampreys is thought to have separated from the lineage leading to the other vertebrates before the jaw arose. Therefore, we have included the lamprey as the outgroup for our analysis. Because derived traits are traits acquired by other members of the vertebrate lineage *after* they diverged from the outgroup, any trait that is present in both the lamprey and the other vertebrates is judged to be ancestral.

We begin by noting that the chimpanzee and mouse share two derived traits—mammary glands and fur—that are absent in both the outgroup and in the other species of the ingroup. Therefore, we infer that mammary glands and fur are derived traits that evolved in a common ancestor of chimpanzees and mice after that lineage separated from the lineages leading to the other vertebrates. In other words, we provisionally assume that mammary glands and fur evolved only once among the animals in our ingroup. These characters are synapomorphies that unite chimpanzees and mice (as well as all other mammals, although we have not included other mammalian species in this example). By the same reasoning, we can infer that the other shared derived traits are synapomorphies for the various groups in which they are expressed. For instance, keratinous scales are a synapomorphy of the lizard, crocodile, and pigeon.

Table 22.1 also tells us that, among the animals in our ingroup, the pigeon has a unique trait: the presence of feathers. Feathers are a synapomorphy of birds, but since we only have one bird in this example, the presence of feathers provides no clues concerning relationships among the eight species of vertebrates we have sampled. However, gizzards are found in birds and crocodiles, so this trait is evidence of a close relationship between birds and crocodilians.

By combining information about the various synapomorphies, we can construct a phylogenetic tree. We infer, for example, that mice and chimpanzees, the only two animals that share fur and mammary glands in our example, share a more recent common ancestor with each other than they do with

TABLE 22.1
Eight Vertebrates Ordered According to Unique Shared Derived Traits

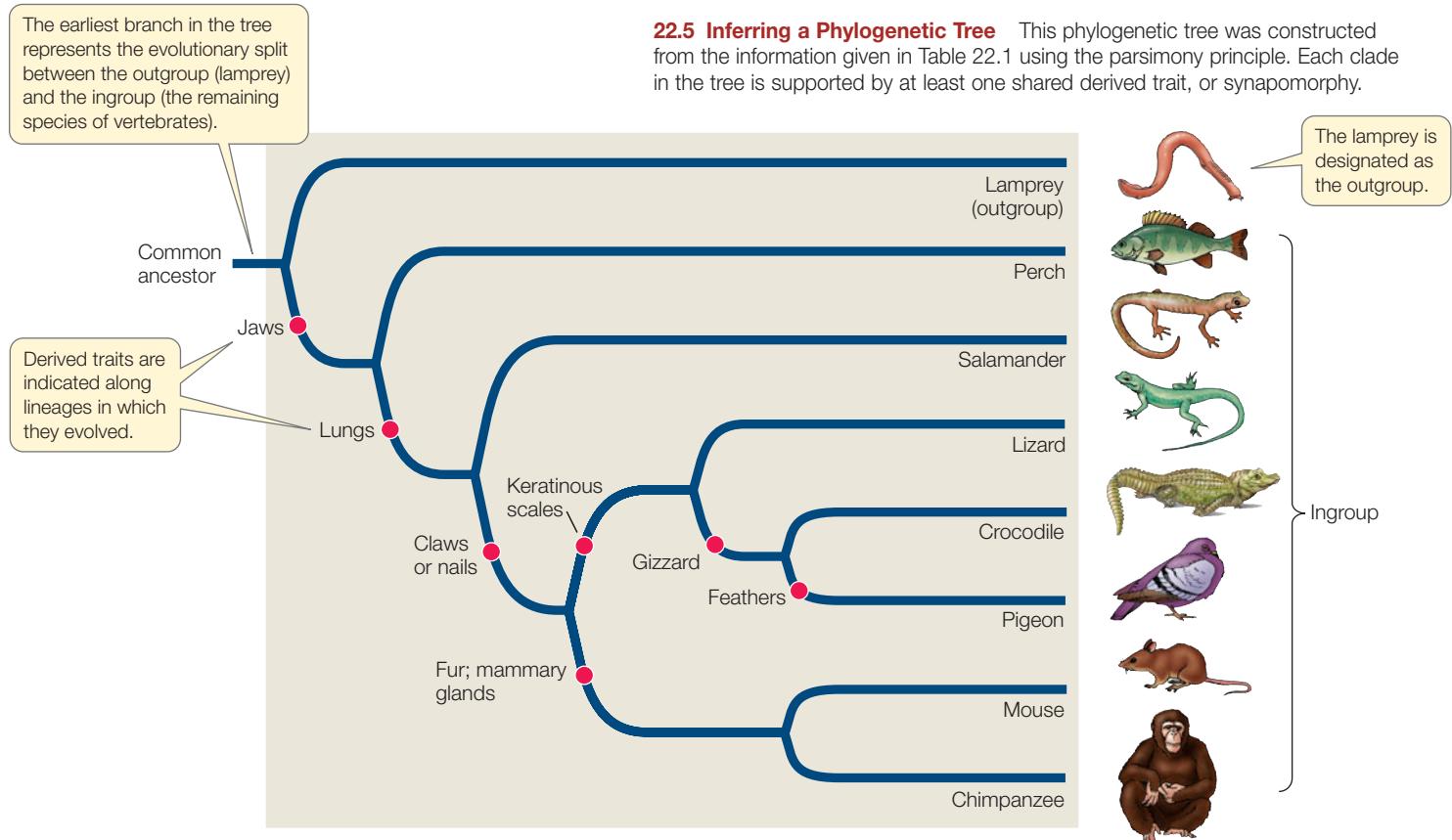
TAXON	DERIVED TRAIT ^a							
	JAWS	LUNGS	CLAWS OR NAILS	GIZZARD	FEATHERS	FUR	MAMMARY GLANDS	KERATINOUS SCALES
Lamprey (outgroup)	-	-	-	-	-	-	-	-
Perch	+	-	-	-	-	-	-	-
Salamander	+	+	-	-	-	-	-	-
Lizard	+	+	+	-	-	-	-	+
Crocodile	+	+	+	+	-	-	-	+
Pigeon	+	+	+	+	+	-	-	+
Mouse	+	+	+	-	-	+	+	-
Chimpanzee	+	+	+	-	-	+	+	-

^aA plus sign indicates the trait is present, a minus sign that it is absent.

pigeons and crocodiles. Otherwise, we would need to assume that the ancestors of pigeons and crocodiles also had fur and mammary glands but subsequently lost them—unnecessary additional assumptions.

Figure 22.5 shows a phylogenetic tree for the vertebrates in Table 22.1, based on the shared derived traits we examined and the assumption that each derived trait evolved only once. This particular tree was easy to construct because the animals and

characters we chose met the assumptions that derived traits appeared only once and were never lost after they appeared. Had we included a snake in the group, our second assumption would have been violated, because we know that the lizard ancestors of snakes had limbs that were subsequently lost. We would need to examine additional characters to determine that the lineage leading to snakes separated from the one leading to lizards long after the lineage leading to lizards separated from



the others. In fact, the analysis of several characters shows that snakes evolved from burrowing lizards that became adapted to a subterranean existence.

Parsimony provides the simplest explanation for phylogenetic data

The phylogenetic tree shown in Figure 22.5 is based on only a very small sample of traits. Typically, biologists construct phylogenetic trees using hundreds or thousands of traits. With larger data sets, we would expect to observe some traits that have changed more than once, and thus we would expect to see some convergence and evolutionary reversal. How do we determine which traits are synapomorphies and which are homoplasies? One way is to invoke the principle of *parsimony*.

In its most general form, the **parsimony principle** states that the preferred explanation of observed data is the simplest explanation. Applying the principle of parsimony to the reconstruction of phylogenies entails minimizing the number of evolutionary changes that need to be assumed over all characters in all groups in the tree. In other words, the best hypothesis under the parsimony principle is one that requires the fewest homoplasies. This application of parsimony is a specific case of a general principle of logic called *Occam's razor*: the best explanation is the one that fits the data best while making the fewest assumptions.

We apply the parsimony principle in constructing phylogenetic trees not because all evolutionary changes always occurred parsimoniously, but because it is logical to adopt the simplest explanation that can account for the observed data. More complicated explanations are accepted only when the evidence requires them. Phylogenetic trees represent our best estimates about evolutionary relationships. They are continually modified as additional evidence becomes available.

Phylogenies are reconstructed from many sources of data

Naturalists have constructed various forms of phylogenetic trees for more than 150 years. In fact, the only figure in the first edition of Darwin's *Origin of Species* was a phylogenetic tree. Tree construction has been revolutionized, however, by the advent of computer software for trait analysis and tree construction, allowing us to consider far more data than could ever before be processed. Combining this with the massive comparative data sets being generated through studies of genomes, biologists are learning details about the tree of life at a remarkable pace.

Any trait that is genetically determined, and therefore heritable, can be used in a phylogenetic analysis. Evolutionary relationships can be revealed through studies of morphology, development, the fossil record, behavioral traits, and molecular traits such as DNA and protein sequences. Let's take a closer look at the types of data used in modern phylogenetic analyses.

yourBioPortal.com
GO TO Web Activity 22.1 • Constructing a
Phylogenetic Tree

MORPHOLOGY An important source of phylogenetic information is *morphology*: the presence, size, shape, and other attributes of body parts. Since living organisms have been observed, depicted, and studied for millenia, we have a wealth of recorded morphological data as well as extensive museum and herbarium collections of organisms whose traits can be measured. New technological tools, such as the electron microscope and computed tomography (CT) scans, enable systematists to examine and analyze the structures of organisms at much finer scales than was formerly possible.

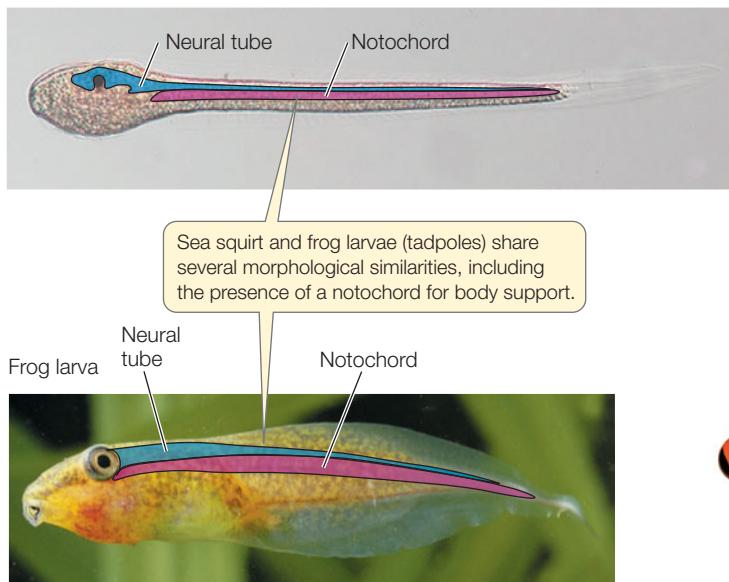
Most species are described and known primarily by their morphology, and morphology provides the most comprehensive data set available for many taxa. The features of morphology that are important for phylogenetic analysis are often specific to a particular group of organisms. For example, the presence, development, shape, and size of various features of the skeletal system are important for the study of vertebrate phylogeny, whereas floral structures are important for studying the relationships among flowering plants (*angiosperms*).

Although often useful, morphological approaches to phylogenetic analysis have some limitations. Some taxa exhibit little morphological diversity, despite great species diversity. For example, the phylogeny of the leopard frogs of North and Central America would be difficult to infer from morphological differences alone, because the many species look very similar, despite important differences in their behavior and physiology. At the other extreme, few morphological traits can be compared across distantly related species (consider earthworms and mammals, for instance). Some morphological variation has an environmental (rather than a genetic) basis and so must be excluded from phylogenetic analyses. An accurate phylogenetic analysis often requires information beyond that supplied by morphology.

DEVELOPMENT Observations of similarities in developmental patterns may reveal evolutionary relationships. Some organisms exhibit similarities in early developmental stages only. The larvae of marine creatures called sea squirts, for example, have a flexible gelatinous rod in the back—the *notochord*—that disappears as the larvae develop into adults. All vertebrate animals also have a notochord at some time during their development (Figure 22.6). This shared structure is one of the reasons for inferring that sea squirts are more closely related to vertebrates than would be suspected if only adult sea squirts were examined.

PALEONTOLOGY The fossil record is another important source of information on evolutionary history. Fossils show us where and when organisms lived in the past and give us an idea of what they looked like. Fossils provide important evidence that helps us distinguish ancestral from derived traits. The fossil record can also reveal when lineages diverged and began their independent evolutionary histories. Furthermore, in groups with few species that have survived to the present, information on extinct species is often critical to an understanding of the large divergences among the surviving species. The fossil record does have limitations, however. Few or no fossils have been found for some groups, and the fossil record for many groups is fragmentary.

Sea squirt larva



Adult



Despite the similarity of their larvae, the morphology of adult frogs and sea squirts provides little evidence of the common ancestry of these two groups.

22.6 The Evolutionary Relationship Between Sea Squirts and Vertebrates

All chordates—a taxonomic group that includes sea squirts and frogs—have notochords at some stage of their development. The larvae share similarities that are not apparent in the adults. Such similarities in development can provide useful evidence of evolutionary relationships. The notochord is lost in adult sea squirts. In adult frogs, as in all vertebrates, the vertebral column replaces the notochord as the support structure.

BEHAVIOR Some behavioral traits are culturally transmitted and some are inherited. If a particular behavior is culturally transmitted, it may not accurately reflect evolutionary relationships (but may nonetheless reflect cultural connections). Bird songs, for instance, are often learned and may be inappropriate traits for phylogenetic analysis. Frog calls, however, are genetically determined and appear to be acceptable sources of information for reconstructing phylogenies.

MOLECULAR DATA All heritable variation is encoded in DNA, and so the complete genome of an organism contains an enormous set of traits (the individual nucleotide bases of DNA) that can be used in phylogenetic analyses. In recent years, DNA sequences have become among the most widely used sources of data for constructing phylogenetic trees. Comparisons of nucleotide sequences are not limited to the DNA in the cell nucleus. Eukaryotes have genes in their mitochondria as well as in their nuclei; plant cells also have genes in their chloroplasts. The chloroplast genome (cpDNA), which is used extensively in phylogenetic studies of plants, has changed slowly over evolutionary time, so it is often used to study relatively ancient phylogenetic relationships. Most animal mitochondrial DNA (mtDNA) has changed more rapidly, so mitochondrial genes have been used extensively to study evolutionary relationships among closely related animal species (the mitochondrial genes of plants evolve more slowly). Many nuclear gene sequences are also commonly analyzed, and now that several entire genomes have been sequenced, they too are used to construct phylogenetic trees. Information on gene products (such as the amino acid sequences of proteins) is also widely used for phylogenetic analyses, as we discuss in Chapter 24.

Mathematical models expand the power of phylogenetic reconstruction

As biologists began to use DNA sequences to infer phylogenies in the 1970s and 1980s, they developed explicit mathematical models describing how DNA sequences change over time. These models account for multiple changes at a given position in a DNA sequence. They also take into account different rates of change at different positions in a gene, at different positions in a codon, and among different nucleotides (see Section 24.1). For example, *transitions* (changes between two purines or between two pyrimidines) are usually more likely than are *transversions* (changes between a purine and pyrimidine).

Mathematical models can be used to compute how a tree might evolve given the observed data. A **maximum likelihood** method will identify the tree that most likely produced the observed data, given the assumed model of evolutionary change. Maximum likelihood methods can be used for any kind of characters, but they are most often used with molecular data, for which explicit mathematical models of evolutionary change are easier to develop. The principal advantages to maximum likelihood analyses are that they incorporate more information about evolutionary change than do parsimony methods, and they are easier to treat in a statistical framework. The principal disadvantages are that they are computationally intensive and require explicit models of evolutionary change (which may not be available for some kinds of character change).

The accuracy of phylogenetic methods can be tested

If phylogenetic trees represent reconstructions of past events, and if many of these events occurred before any humans were

around to witness them, how can we test the accuracy of phylogenetic methods? Biologists have conducted experiments both in living organisms and with computer simulations that have demonstrated the effectiveness and accuracy of phylogenetic methods.

In one experiment designed to test the accuracy of phylogenetic analysis, a single viral culture of bacteriophage T7 was used as a starting point, and lineages were allowed to evolve from this ancestral virus in the laboratory (Figure 22.7). The initial culture was split into two separate lineages, one of which became the ingroup for analysis and the other of which became the outgroup for rooting the tree. The lineages in the ingroup were split in two after every 400 generations, and samples of the virus were saved for analysis at each branching point. The lineages were allowed to evolve until there were eight lineages in the ingroup. Mutagens were added to the viral cultures to increase the mutation rate so that the amount of change and the degree of homoplasy would be typical of the organisms analyzed in average phylogenetic analyses. The investigators then sequenced samples from the end points of the eight lineages, as well as from the ancestors at the branching points. They then gave the sequences from the end points of the lineages to other investigators to analyze, without revealing the known history of the lineages or the sequences of the ancestral viruses.

After the phylogenetic analysis was completed, the investigators asked two questions: Did phylogenetic methods reconstruct the known history correctly, and were the sequences of the ancestral viruses reconstructed accurately? The answer in both cases was yes: the branching order of the lineages was reconstructed exactly as it had occurred, more than 98 percent of the nucleotide positions of the ancestral viruses were reconstructed correctly, and 100 percent of the amino acid changes in the viral proteins were reconstructed correctly.

[yourBioPortal.com](#)

GO TO Animated Tutorial 22.1 • Using Phylogenetic Analysis to Reconstruct Evolutionary History

The experiment shown in Figure 22.7 demonstrated that phylogenetic analysis was accurate under the conditions tested, but it did not examine all possible conditions. Other experimental studies have taken other factors into account, such as the sensitivity of phylogenetic analysis to convergent environments and highly variable rates of evolutionary change. In addition, computer simulations based on evolutionary models have been used extensively to study the effectiveness of phylogenetic analysis. These studies have also confirmed the accuracy of phylogenetic methods and have been used to refine those methods and extend them to new applications.

INVESTIGATING LIFE

22.7 The Accuracy of Phylogenetic Analysis

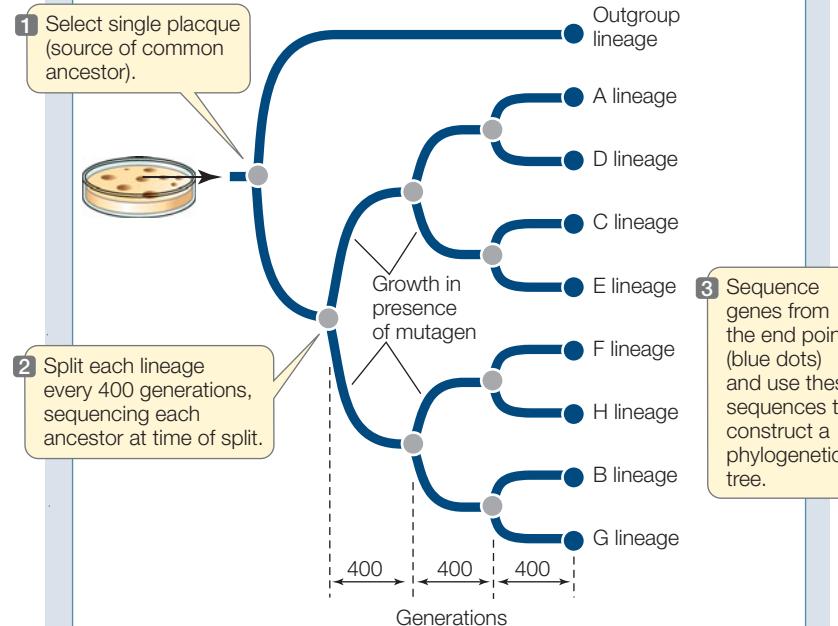
To test whether analysis of gene sequences can accurately reconstruct evolutionary phylogeny, we must have an unambiguously known phylogeny to compare against the reconstruction. Will the observed phylogeny match the reconstruction?

HYPOTHESIS

A phylogeny reconstructed from analysis of the DNA sequences of living organisms can accurately match the known evolutionary history of the organisms.

METHOD

In the laboratory, researchers produced an unambiguous phylogeny of nine viral lineages, enhancing the mutation rate to increase variation among the lineages.



Viral sequences from the end points of each lineage (blue dots) were subjected to phylogenetic analysis by investigators who were unaware of the history of the lineages or the gene sequences of the ancestral viruses. These investigators reconstructed the phylogeny based solely on their analyses of the descendants' genomes.

RESULTS

The true phylogeny and ancestral DNA sequences were accurately reconstructed solely from the DNA sequences of the viruses at the tips of the tree.

CONCLUSION

Phylogenetic analysis of DNA sequences can accurately reconstruct evolutionary history.

FURTHER INVESTIGATION: The lineages in this experiment evolved under similar conditions. How might changing environmental conditions for some of the lineages affect the result?

Go to [yourBioPortal.com](#) for original citations, discussions, and relevant links for all INVESTIGATING LIFE figures.

22.2 RECAP

Phylogenetic trees can be constructed by using the parsimony principle to find the simplest explanation for the evolution of traits. Maximum likelihood methods incorporate more explicit models of evolutionary change to reconstruct evolutionary history.

- Do you understand how a phylogenetic tree is constructed? See pp. 468–471 and Figure 22.5
- Is there a way to test whether phylogenetic trees provide accurate reconstructions of evolutionary history? See p. 474 and Figure 22.7

Biologists in many fields now routinely reconstruct phylogenetic relationships. Let's examine some of the many uses of these phylogenetic trees.

22.3 How Do Biologists Use Phylogenetic Trees?

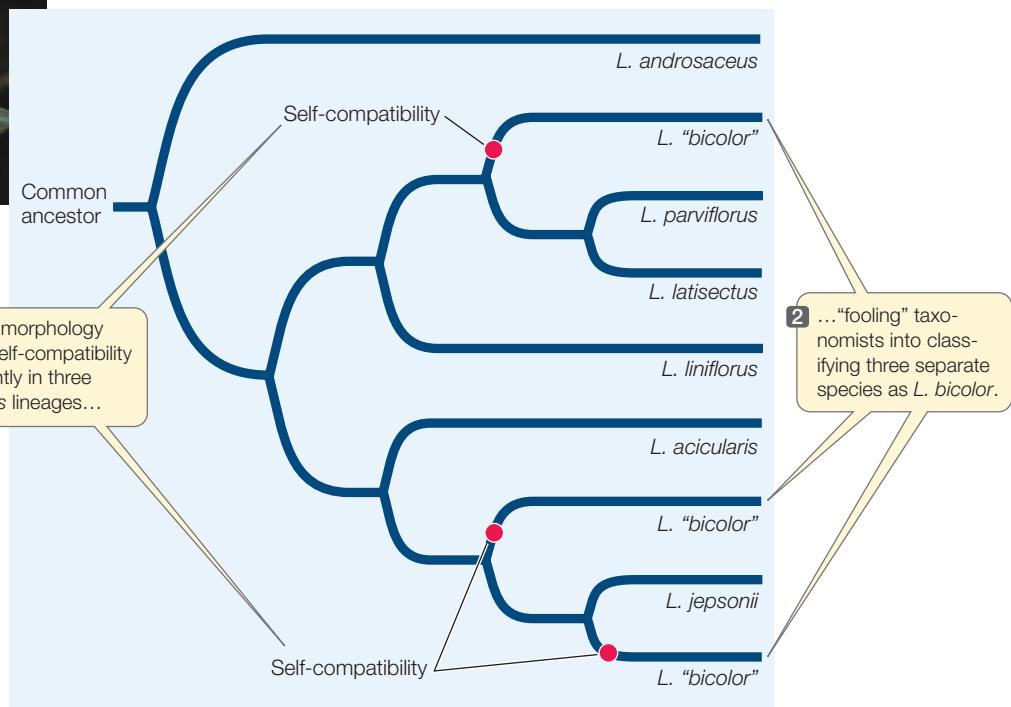
Information about the evolutionary relationships among organisms is useful to scientists investigating a wide variety of biological questions. In this section we will illustrate how phylogenetic trees can be used to ask questions about the past, and to compare aspects of the biology of organisms in the present.



1 Convergent floral morphology associated with self-compatibility arose independently in three different *Linanthus* lineages...

22.8 Phylogeny of a Section of the Plant Genus *Linanthus*

Self-compatibility apparently evolved independently three times in this group. Because the appearance and structure of the flowers converged in the three selfing lineages, taxonomists mistakenly thought they were varieties of the same species.



Phylogenies help us reconstruct the past

Most flowering plants reproduce by mating with another individual—a process called *outcrossing*. Many outcrossing species have mechanisms to prevent self-fertilization, and so are referred to as *self-incompatible*. Individuals of some species, however, regularly fertilize themselves with their own pollen; they are termed *selfing* species, which of course requires that they be *self-compatible*. How can we tell how often self-compatibility has evolved in a group of plants? We can do so by conducting a phylogenetic analysis of outcrossing and selfing species and testing the species for self-compatibility.

The evolution of fertilization mechanisms was examined in *Linanthus*, a genus in the phlox family that exhibits a diversity of breeding systems and pollination mechanisms. The outcrossing species of *Linanthus* have long petals, are pollinated by long-tongued flies, and are self-incompatible. The self-pollinating species of *Linanthus*, in contrast, all have short petals and do not require insect pollinators to reproduce successfully. Investigators reconstructed a phylogeny for 12 species in the genus using nuclear ribosomal DNA sequences (Figure 22.8). They determined whether each species was self-compatible by artificially pollinating flowers with the plant's own pollen or with pollen from other individuals and observing whether viable seeds formed.

Several lines of evidence suggest that self-incompatibility is the ancestral state in *Linanthus*. Multiple origins of self-incompatibility have not been found in any flowering plant family to date. Self-incompatibility depends on physiological mechanisms in both the pollen and the stigma (the female organ on which pollen lands) and is under the control of at least three different alleles. Therefore, a change from self-incompatibility to

22.9 Phylogenetic Tree of Immunodeficiency Viruses

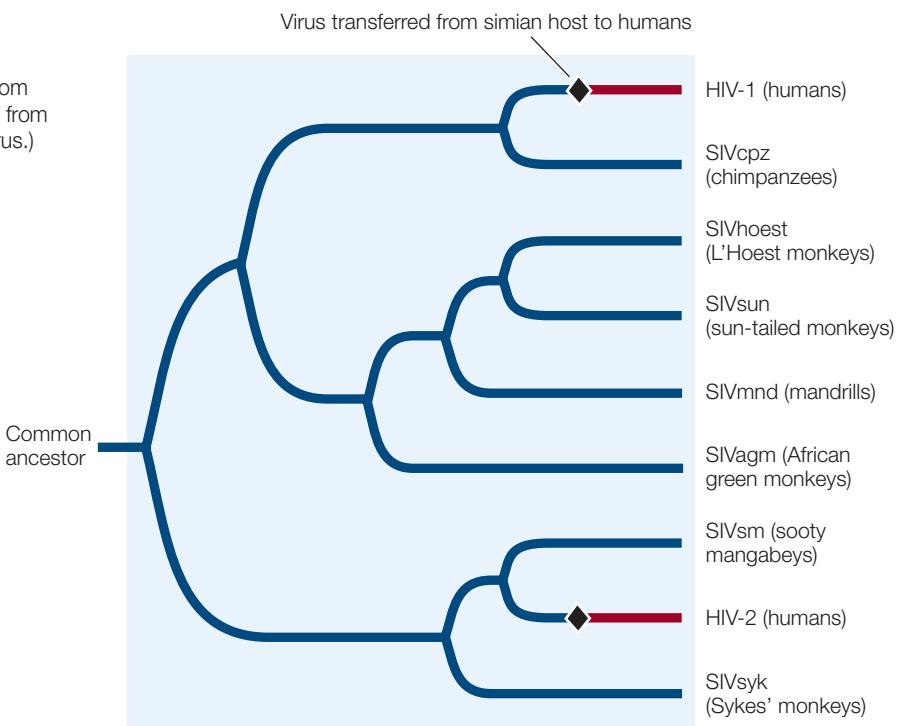
Immunodeficiency viruses have been transmitted to humans from two different simian hosts: HIV-1 from chimpanzees and HIV-2 from sooty mangabeys. (SIV stands for simian immunodeficiency virus.)

self-compatibility would be easier than the reverse change. In addition, in all self-incompatible species of *Linanthus*, the site of pollen rejection is the stigma, even though sites of pollen rejection vary greatly among other plant families.

Assuming that self-incompatibility is the ancestral state, the reconstructed phylogeny suggests that self-compatibility evolved three times within this group of *Linanthus* (see Figure 22.8). The change to self-compatibility has been accompanied by the evolution of reduced petal size. Interestingly, the striking similarity of the flowers in the self-compatible groups once led to their being classified as members of a single species. The phylogenetic analysis using ribosomal DNA showed them to be members of three distinct lineages, however.

Reconstructing the past is important for understanding many biological processes. In the case of zoonotic diseases (diseases caused by infectious organisms transmitted to humans from another animal host), it is important to understand when, where, and how the disease first entered a human population. Human immunodeficiency virus (HIV) is the cause of such a zoonotic disease, acquired immunodeficiency syndrome, or AIDS. As we described in the opening to this chapter, phylogenetic analyses have become important for studying the transmission of viruses such as HIV. Phylogenies are also important for understanding the present global diversity of HIV and for determining the virus's origins in human populations. A broader phylogenetic analysis of immunodeficiency viruses shows that humans acquired these viruses from two different hosts: HIV-1 from chimpanzees, and HIV-2 from sooty mangabeys (Figure 22.9).

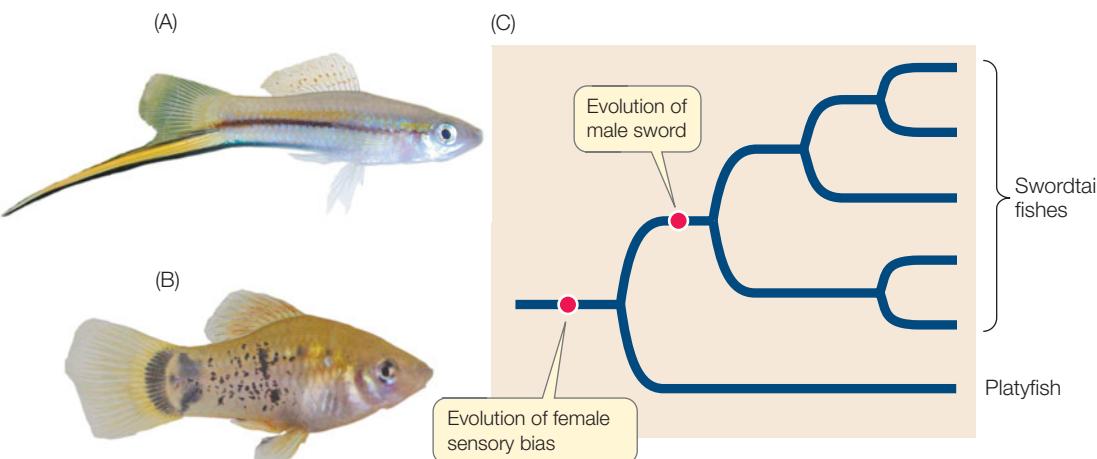
HIV-1 is the common form of the virus in human populations in central Africa, where chimpanzees are hunted for food, and HIV-2 is the common form in human populations in western Africa, where sooty mangabeys are hunted for food. Thus it seems



likely that these viruses entered human populations through hunters who cut themselves while skinning chimpanzees and sooty mangabeys. The relatively recent global pandemic of AIDS occurred when these infections in local African populations rapidly spread through human populations around the world.

Phylogenies allow us to compare and contrast living organisms

Male swordtails—a group of fishes in the genus *Xiphophorus*—have a long, colorful tail extension (Figure 22.10A), and their reproductive success is closely associated with this appendage. Males with a long sword are more likely to mate successfully than are males with a short sword (an example of *sexual selection*; see Chapters 21 and 23). Several explanations have been advanced for the evolution of this structure, including the hypothesis that the sword simply exploits a preexisting bias in the sensory system of the females. This *sensory exploitation hypothesis*



22.10 The Origin of a Sexually Selected Trait

(A) The large tail of male swordtail fishes (genus *Xiphophorus*) apparently evolved through sexual selection, with females mating preferentially with males with a longer "sword." (B) A male platyfish, member of a related species. (C) Phylogenetic analysis reveals that the platyfishes split from the swordtails before the evolution of the sword. The independent finding that female platyfishes prefer males with an artificial sword further supports the idea that this appendage evolved as a result of a preexisting preference in the females.

esis suggests that female swordtails had a preference for males with long tails even before the tails evolved (perhaps because females assess the size of males by their total body length—including the tail—and prefer larger males).

To test the sensory exploitation hypothesis, a phylogeny was used to identify the swordtail relatives that had split most recently from their lineage before the evolution of sword extensions. These closest relatives turned out to be the platyfishes, another group of *Xiphophorus* (Figure 22.10B). Even though male platyfishes do not normally have swords, when researchers attached artificial swordlike structures to the tails of some male platyfishes, female platyfishes preferred the males with an artificial sword, thus providing support for the hypothesis that female *Xiphophorus* had a preexisting sensory bias favoring tail extensions even before the trait evolved (Figure 22.10C). Thus, a long tail became a sexually selected trait because of the preexisting preference of the females.

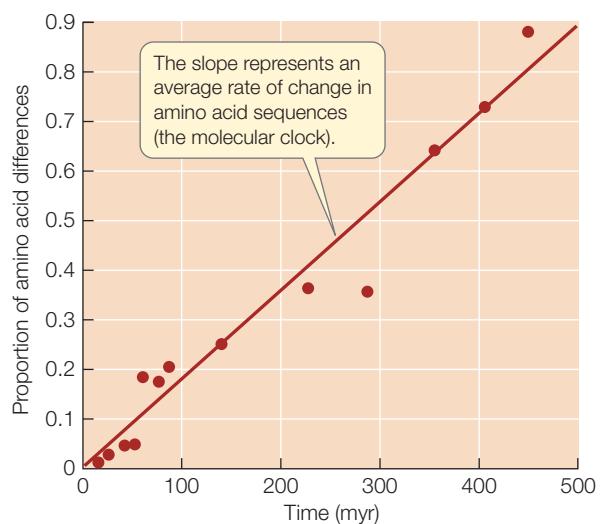
Ancestral states can be reconstructed

In addition to using phylogenetic methods to infer evolutionary relationships among lineages, biologists can use them to reconstruct the morphology, behavior, or nucleotide and amino acid sequences of ancestral species (as was demonstrated for the ancestral sequences of bacteriophage T7 in the experiment shown in Figure 22.7). For instance, a phylogenetic analysis was used to reconstruct an opsin protein in the ancestral archosaur (the most recent common ancestor of birds, dinosaurs, and crocodiles). Opsins are pigment proteins involved in vision; different opsins (with different amino acid sequences) are excited by different wavelengths of light. Knowledge of the opsin sequence in the ancestral archosaur would provide clues about the animal's visual capabilities and therefore about some of its probable behaviors. Investigators used phylogenetic analysis of opsin from living vertebrates to estimate the amino acid sequence of the pigment that existed in the ancestral archosaur. A protein with this same sequence was then constructed in the laboratory. The investigators tested the reconstructed opsin and found a significant shift toward the red end of the spectrum in the light sensitivity of this protein compared with that of most modern opsins. Modern species that exhibit similar sensitivity are adapted for nocturnal vision, so the investigators inferred that the ancestral archosaur might have been active at night. Thus, reminiscent of the movie *Jurassic Park*, phylogenetic analyses are being used to reconstruct extinct species, one protein at a time.

Molecular clocks help date evolutionary events

For many applications, biologists want to know not only the order in which evolutionary lineages split but also the timing of those splits. In 1965, Emile Zuckerkandl and Linus Pauling hypothesized that rates of molecular change were constant enough that they could be used to predict evolutionary divergence times—an idea that has become known as the *molecular clock hypothesis*.

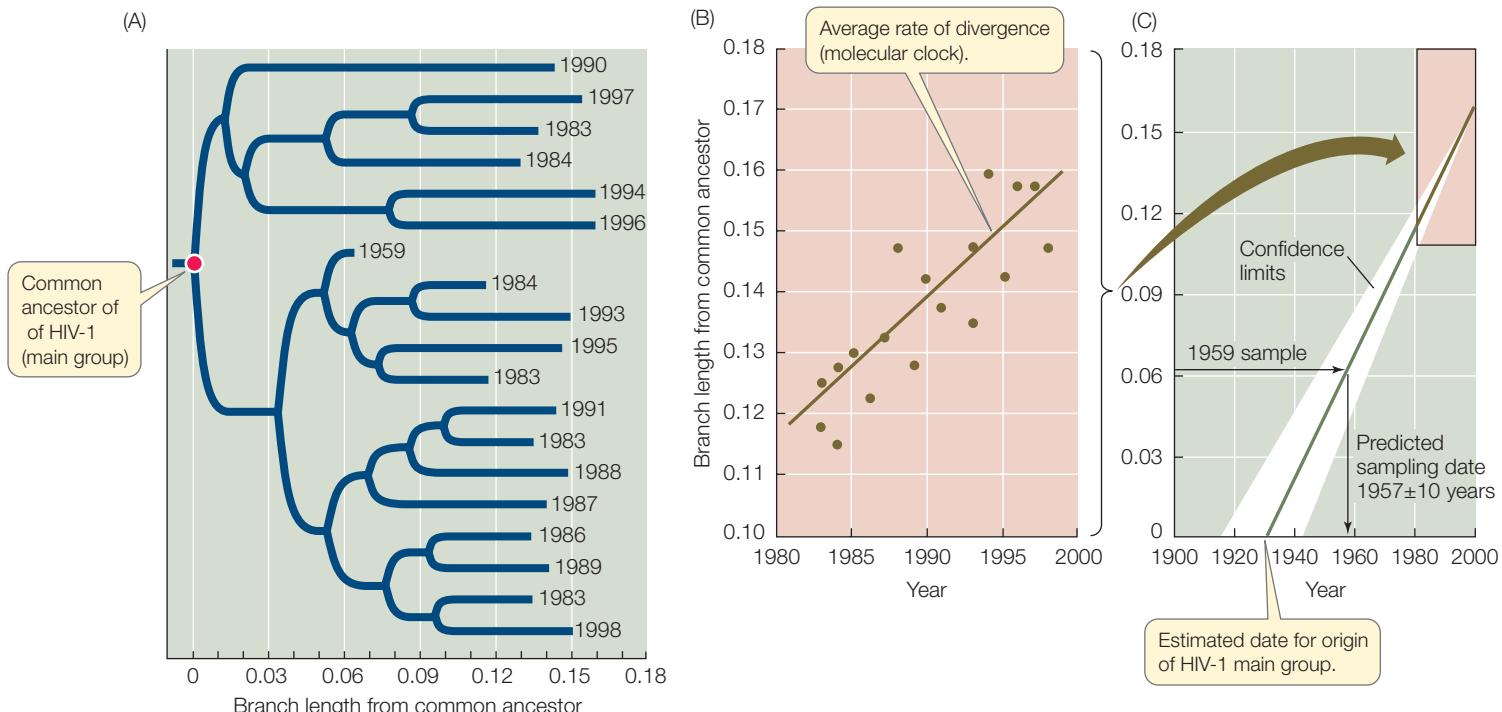
Of course, different genes evolve at different rates, and there are also differences in evolutionary rates among species related



22.11 A Molecular Clock of the Protein Hemoglobin Amino acid replacements in hemoglobin have occurred at a relatively constant rate over nearly 500 million years of evolution. The graph shows the relationship between time of divergence and proportion of amino acid change for 13 pairs of vertebrate hemoglobin proteins. The average rate of change represents the molecular clock for hemoglobin in vertebrates.

to differing generation times, environments, efficiencies of DNA repair systems, and other biological factors. Nonetheless, among closely related species, a given gene usually evolves at a reasonably constant rate. Therefore, the protein encoded by the gene also accumulates amino acid substitutions at a relatively constant rate (Figure 22.11). A **molecular clock** uses the average rate at which a given gene or protein accumulates changes to gauge the time of divergence for a particular split in the phylogeny. Molecular clocks must be calibrated using independent data, such as the fossil record, known times of divergence, or biogeographic dates (such as the dates for separations of continents). Using such calibrations, times of divergence have been estimated for many groups of species that have diverged over millions of years.

Molecular clocks are not only used to date ancient events; they are also used to study the timing of comparatively recent events. Most samples of HIV-1 have been collected from humans only since the early 1980s, although a few isolates from medical biopsies are available from as early as the 1950s. But biologists can use the observed changes in HIV-1 over the past several decades to project back to the common ancestor of all HIV-1 isolates, and estimate when HIV-1 first entered human populations from chimpanzees. The clock can be calibrated using the samples from the 1980s and 1990s, and then tested using the samples from the 1950s. As shown in Figure 22.12C, a sample from a 1959 biopsy is dated by molecular clock analysis at 1957 ± 10 years. The molecular clock was also used to project back to the common ancestor of this group of HIV-1 samples. Extrapolation suggests a date of origin for this group of viruses of about 1930. Although AIDS was unknown to Western medicine until the 1980s, this analysis shows that HIV-1 was present (probably at very low frequency) in human populations in Africa for at least a half-century before its emergence as a



22.12 Dating the Origin of HIV-1 in Human Populations

(A) A phylogenetic analysis of the main group of HIV-1 viruses. The dates indicate the years in which samples were taken. (For clarity, only a small fraction of the samples that were examined in the original study are shown.) (B) A plot of year of isolation versus genetic divergence from the common ancestor provides an average rate of divergence, or a molecular clock. (C) The molecular clock is used to date a sample taken in 1959 (as a test of the clock) and the unknown date of origin of the HIV-1 main group (about 1930).

global pandemic (Figure 22.12). Biologists have used similar analyses to conclude that immunodeficiency viruses have been transmitted repeatedly into human populations from multiple primates for more than a century (see also Figure 22.9).

22.3 RECAP

Phylogenetic trees are used to reconstruct the past history of lineages, to determine when and where traits arose, and to make relevant biological comparisons among genes, populations, and species. They can also be used to reconstruct ancestral traits and to estimate the timing of evolutionary events.

- Explain how phylogenetic trees can help determine the number of times a particular trait evolved. See pp. 473–474 and Figure 22.8
- How does the reconstruction of ancestral traits help biologists explain the evolution of visual pigment proteins? See p. 475
- How do molecular clocks add a time dimension to phylogenetic trees? See p. 475 and Figure 22.12

All of life is connected through evolutionary history, and the relationships among organisms provide a natural basis for making biological comparisons. For these reasons, biologists use phylogenetic relationships as the basis for organizing life into a coherent classification system, described in the next section.

22.4 How Does Phylogeny Relate to Classification?

The biological classification system in widespread use today is derived from a system developed by the Swedish biologist Carolus Linnaeus in the mid-1700s. Linnaeus developed a naming system called **binomial nomenclature** that has allowed scientists throughout the world to refer unambiguously to the same organisms by the same names (Figure 22.13).

Linnaeus gave each species two names, one identifying the species itself and the other the genus to which it belongs. A **genus** (plural, *genera*) is a group of closely related species. Optionally, the name of the taxonomist who first proposed the species name may be added at the end. Thus *Homo sapiens* Linnaeus is the name of the modern human species. *Homo* is the genus to which the species belongs, and *sapiens* identifies the particular species in the genus *Homo*; Linnaeus proposed the species name *Homo sapiens*. You can think of the generic name *Homo* as equivalent to your surname and the specific name *sapiens* as equivalent to your first name. The name of the genus is always capitalized, and the name identifying the species is always lowercased. Both names are italicized, whereas common names of organisms are not. Rather than repeating the name of a genus when it is used several times in the same discussion, biologists often spell it out only once and abbreviate it to the initial letter thereafter (*D. melanogaster* rather than *Drosophila melanogaster*, for example).

(A) *Campanula rotundifolia*(B) *Endymion non-scriptus*(C) *Mertensia virginica*

22.13 Many Different Plants Are Called Bluebells All three of these distantly related plant species are called “bluebells.” Binomial nomenclature allows us to communicate exactly what is being described. (A) *Campanula rotundifolia*, found on the North American Great Plains, belongs to a larger group of bellflowers. (B) *Endymion non-scriptus*, English bluebell, is related to hyacinths. (C) *Mertensia virginica*, Virginia bluebell, belongs in a very different group of plants known as borage.

As we noted earlier, any group of organisms that is treated as a unit in a biological classification system, such as the genus *Drosophila*, or all insects, is called a *taxon*. In the Linnaean system, species and genera are further grouped into a hierarchical system of higher taxonomic categories. The taxon above the genus in the Linnaean system is the **family**. The names of animal families end in the suffix “-idae.” Thus Formicidae is the family that contains all ant species, and the family Hominidae contains humans and our recent fossil relatives, as well as our closest living relatives, the chimpanzees and gorillas. Family names are based on the name of a member genus; Formicidae is based on the genus *Formica*, and Hominidae is based on *Homo*. The same rules are used in classifying plants, except that the suffix “-aceae” is used for plant family names instead of “-idae.” Thus Rosaceae is the family that includes the genus of roses (*Rosa*) and its close relatives. In the Linnaean system, families

are grouped into **orders**, orders into **classes**, and classes into **phyla** (singular *phylum*), and phyla into **kingdoms**. However, Linnaean classification is often subjective; whether a particular taxon is considered, say, an order or a class is often a subjective decision. Today, Linnaean terms are used largely for convenience. Although families are always grouped within orders, orders within classes, and so forth, there is nothing that makes a “family” in one group equivalent (in number of genera or in evolutionary age, for instance) to a “family” in another group.

Linnaeus recognized the overarching hierarchy of life, but he developed his system before evolutionary thought had become widespread. Biologists today recognize the tree of life as the basis for biological classification and often name clades without placing them into any Linnaean rank. But regardless of whether they rank organisms into Linnaean categories or use unranked clade names, modern biologists use evolutionary relationships as the basis for distinguishing biological taxa.

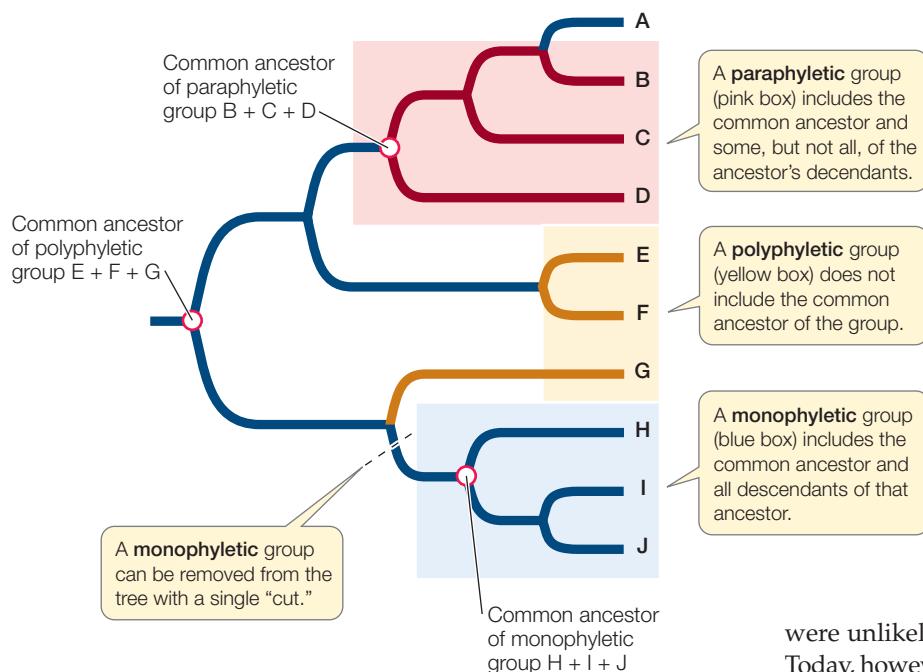
Evolutionary history is the basis for modern biological classification

Biological classification systems are used to express relationships among organisms. The kind of relationship we wish to express influences which features we use to classify organisms. If, for instance, we were interested in a system that would help us decide what plants and animals were desirable as food, we might devise a classification based on tastiness, ease of capture, and the number of edible parts each organism possessed. Early Hindu classifications of organisms were designed according to these criteria. Such systems served the needs of the people who developed them, but are not adequate for formal scientific classification.

Taxonomists today use biological classifications to express the evolutionary relationships of organisms. Taxa are expected to be **monophyletic**, meaning that the taxon contains an ancestor and all descendants of that ancestor, and no other organisms (Figure 22.14). In other words, the taxon is an historical group of related species, or a complete branch on the tree of life (a *clade*). Although biologists seek to describe and name only monophyletic taxa, the detailed phylogenetic information needed to do so is not always available. A group that does not include its common ancestor is called a **polyphyletic** group. A group that does not include all the descendants of a common ancestor is called a **paraphyletic** group.

A true monophyletic group (i.e., a clade) can be removed from a phylogenetic tree by a single “cut” in the tree, as shown in Figure 22.14. Note that there are many monophyletic groups on any phylogenetic tree, and that these groups are successively smaller subsets of larger monophyletic groups. This hierarchy of biological taxa, with all of life as the most inclusive taxon and many smaller taxa within larger taxa, down to the individual species, is the modern basis for biological classification.

Virtually all taxonomists now agree that polyphyletic and paraphyletic groups are inappropriate as taxonomic units, because they do not correctly reflect evolutionary history. The classifications used today still contain such groups because some organisms have not been evaluated phylogenetically. As mistakes in prior classifications are detected, taxonomic names



22.14 Monophyletic, Polyphyletic, and Paraphyletic Groups

Monophyletic groups are the basis of biological taxa in modern classifications. Polyphyletic and paraphyletic groups do not accurately reflect evolutionary history.

yourBioPortal.com

GO TO Web Activity 22.2 • Types of Taxa

are revised and polyphyletic and paraphyletic groups are eliminated from the classifications.

Several codes of biological nomenclature govern the use of scientific names

Several sets of explicit rules govern the use of scientific names. Biologists around the world follow these rules voluntarily to facilitate communication and dialogue. Although there may be dozens of common names for an organism in many different languages, the rules of biological nomenclature are designed so that there is only one correct scientific name for any single recognized taxon and (ideally) a given scientific name applies only to a single taxon (that is, each scientific name is unique). Sometimes the same species is named more than once (when more than one taxonomist has taken up the task); the rules specify that the valid

name is the first name that was proposed. If the same name is inadvertently given to two different species, then a replacement name must be given to the species that was named second.

Because of the historical separation of the fields of zoology, botany (including, originally, the study of fungi), and microbiology, different sets of taxonomic rules were developed for each of these groups. Yet another set of rules for classifying viruses emerged later. This has resulted in many duplicated names in groups that are governed by different sets of rules: *Drosophila*, for instance, is both a genus of fruit flies and a genus of fungi, and there are species in both groups that have identical names. Until recently these duplicated names caused little confusion, since traditionally biologists who studied fruit flies

were unlikely to read the literature on fungi (and vice versa). Today, however, given the use of large, universal biological databases (such as GenBank, which includes DNA sequences from across all life), it is increasingly important that each taxon have a unique name. Taxonomists are now working to develop common sets of rules that can be applied across all living organisms.

22.4 RECAP

Biologists organize and classify life by identifying and naming monophyletic groups. Several sets of rules govern the use of scientific names so that each species and higher taxon can be identified and named unambiguously.

- Explain the difference between monophyletic, paraphyletic, and polyphyletic groups. See p. 477 and Figure 22.14
- Do you understand why biologists prefer monophyletic groups in formal classifications? See p. 477

Now that we have seen how evolution occurs and how phylogenies can be used to study evolutionary relationships, we are ready to consider the process of speciation. Speciation is what leads to the splitting events on the tree of life, and is the process that results in the millions of species that constitute biodiversity.

CHAPTER SUMMARY

22.1 What Is Phylogeny?

- Phylogeny** is the history of descent of organisms from their common ancestor. Groups of evolutionarily related species are represented as related branches in a **phylogenetic tree**.

Review Figure 22.2

- A group of species that consists of all the evolutionary descendants of a common ancestor is called a **clade**. Named clades and species are called **taxa**.
- Homologies** are similar traits that have been inherited from a common ancestor. Review Figure 22.4

- A trait that is shared by two or more taxa and is derived through evolution from a common ancestral form is called a **synapomorphy**.
- Similar traits may occur among species that do not result from common ancestry. **Convergent evolution** and **evolutionary reversals** can give rise to such traits, which are called **homoplasies**.

22.2 How Are Phylogenetic Trees Constructed?

SEE WEB ACTIVITY 22.1

- Phylogenetic trees can be inferred from synapomorphies using the principle of **parsimony**. **Review Figure 22.5**
- Sources of phylogenetic information include morphology, patterns of development, the fossil record, behavioral traits, and molecular traits such as DNA and protein sequences.
- Phylogenetic trees can also be inferred with **maximum likelihood** methods, which calculate the probability that a particular tree will have generated the observed data.

22.3 How Do Biologists Use Phylogenetic Trees?

- Phylogenetic trees are used to reconstruct the past and understand the origin of traits. **Review Figure 22.8**
- Phylogenetic trees are used to make appropriate evolutionary comparisons among living organisms.
- Biologists can use phylogenetic trees to reconstruct ancestral states. **SEE ANIMATED TUTORIAL 22.1**
- Phylogenetic trees may include estimates of times of divergence of lineages determined by **molecular clock** analysis. **Review Figure 22.12**

22.4 How Does Phylogeny Relate to Classification?

- Taxonomists organize biological diversity on the basis of evolutionary history.
- Taxa in modern classifications are expected to be **monophyletic** groups. **Paraphyletic** and **polyphyletic** groups are not considered appropriate taxonomic units. **Review Figure 22.14, WEB ACTIVITY 22.2**
- Several sets of rules govern the use of scientific names, with the goal of providing unique and universal names for biological taxa.

SELF-QUIZ

1. A *clade* is
 - a type of phylogenetic tree.
 - a group of evolutionarily related species that share a common ancestor.
 - a tool for constructing phylogenetic trees.
 - an extinct species.
 - an ancestral species.
2. Phylogenetic trees may be constructed for
 - genes.
 - species.
 - major evolutionary groups.
 - viruses.
 - All of the above.
3. A shared derived trait, used as the basis for inferring a monophyletic group, is called
 - a synapomorphy.
 - a homoplasy.
 - a parallel trait.
 - a convergent trait.
 - a phylogeny.
4. The parsimony principle can be used to infer phylogenetic trees because
 - evolution is nearly always parsimonious.
 - it is logical to adopt the simplest hypothesis capable of explaining the known facts.
 - once a trait changes, it never reverses condition.
 - all species have an equal probability of evolving.
 - closely related species are always very similar to one another.
5. Convergent evolution and evolutionary reversal are two sources of
 - homology.
 - parsimony.
 - synapomorphy.
 - monophly.
 - homoplasy.
6. Which of the following are commonly used to infer phylogenetic relationships among plants but not among animals?
 - Nuclear genes
 - Chloroplast genes
 - Mitochondrial genes
 - Ribosomal RNA genes
 - Protein-coding genes
7. Which of the following is *not* true of maximum likelihood or parsimony methods for inferring phylogeny?
 - The maximum likelihood method requires an explicit model of evolutionary character change.
 - The parsimony method is computationally easier than the maximum likelihood method.
 - The maximum likelihood method is easier to treat in a statistical framework.
 - The maximum likelihood method is most often used with molecular data.
 - Parsimony is usually used to infer time on a phylogenetic tree.
8. Taxonomists strive to include taxa in biological classifications that are
 - monophyletic.
 - paraphyletic.
 - polyphyletic.
 - homoplastic.
 - monomorphic.
9. Which of the following groups have separate sets of rules for nomenclature?
 - Animals
 - Plants and fungi
 - Bacteria
 - Viruses
 - All of the above

10. If two scientific names are proposed for the same species, how do taxonomists decide which name should be used?
- The name that provides the most accurate description of the organism is used.
 - The name that was proposed most recently is used.
 - The name that was used in the most recent taxonomic revision is used.
 - The first name to be proposed is used, unless that name was previously used for another species.
 - Taxonomists use whichever name they prefer.

FOR DISCUSSION

- Why are taxonomists concerned with identifying species that share a particular common ancestor?
- How are fossils used to identify ancestral and derived traits of organisms? How can fossils be integrated into phylogenetic analyses?
- The parsimony principle is often used to construct phylogenetic trees. What are the limitations of parsimony, and why do some biologists prefer model-based approaches, such as maximum likelihood methods?
- A student of the evolution of frogs has proposed a strikingly new classification of frogs based on an analysis of a few mitochondrial genes from about 10 percent of frog species. Should frog taxonomists immediately accept the new classification? Why or why not?
- What are some of the assumptions that go into a molecular clock analysis? How could these assumptions be violated? How could molecular clock analyses be modified to consider these additional sources of variation?
- Classification systems summarize much information about organisms and enable us to remember the traits of many organisms. From your general knowledge, how many traits can you associate with the following names: conifer, fern, bird, mammal?

ADDITIONAL INVESTIGATION

West Nile virus kills birds of many species and can cause fatal encephalitis (inflammation of the brain) in humans and horses. The virus was first isolated in Africa (where it is thought to be endemic) in the 1930s, and by the 1990s it had been found throughout much of Eurasia. West Nile virus was not found in

North America until 1999, but since that time it has spread rapidly across most of the United States. The genome of West Nile virus evolves quickly. How could you use phylogenetic analysis to investigate the geographic origin of the West Nile virus that was introduced into North America in 1999?

WORKING WITH DATA ([GO TO yourBioPortal.com](#))

Constructing a Phylogenetic Tree In this exercise based on Figure 22.7, you will use a subset of the DNA sequences from the experimental lineages to reconstruct the evolutionary rela-

tionships among the viruses. You will also use these data to reconstruct the DNA sequences of the viral ancestors.

Catching speciation in the act

When biologists first explored the Cuatro Ciénegas basin of northern Mexico, they found many organisms that are not found anywhere else in the world. So far researchers have described about 150 species of plants and animals that are restricted to this small region. Even though Cuatro Ciénegas is in a desert, about 30 of these unique species are aquatic, living in the isolated springs and marshes of the basin. An unusual aquatic box turtle, beautiful cichlid fishes, and tiny crustaceans are among the many aquatic species that are confined to Cuatro Ciénegas. Why are so many different species found here and nowhere else?

Biologists and geologists found that, over the past several million years, this desert oasis has repeatedly been isolated by a succession of geological events that cut it off from the river systems and mountain ranges of northern Mexico. Many different *speciation events* associated with these geological events make Cuatro Ciénegas a natural laboratory for studying speciation by geographic isolation.

Each time gene flow between organisms in the basin and the surrounding areas ceased, populations living inside and outside the basin began to diverge from one another. Over thousands of generations of such isolation, new species developed. These new species no longer share the same gene pool, are adapted to different environments, and look different from one another. And—extremely important—the organisms have diverged to the point that they are no longer capable of reproducing with one another—one of the hallmarks of distinct species.

Although speciation is often studied in natural settings such as Cuatro Ciénegas, some aspects of speciation can be studied in controlled laboratory experiments, using organisms with short generation times. For example, William Rice and George Salt conducted an experiment in which fruit flies were allowed to choose food sources in different habitats, where mating also took place. The habitats were vials in different parts of an experimental cage. The habitats differed in three parameters: (1) light; (2) the direction in which the fruit flies could move (up or down); and (3) concentrations of two aromatic chemicals, ethanol and acetaldehyde. In just 35 generations, two groups of flies were genetically and reproductively isolated from one another because they had evolved distinct preferences for different habitats. In controlled experiments like these, biologists are beginning to study and understand the genetic details of speciation.



A Natural Laboratory A swimmer surveys several of the fish species that are isolated in the desert oasis of the Cuatro Ciénegas basin in northern Mexico.



Experimental Subjects Fruit flies of the genus *Drosophila* are easily reared in the laboratory. Their short generation time (7–10 days from newly laid egg to reproductive adult) makes them ideal subjects for controlled experiments on speciation.

The *origin of species*—the splitting and diverging of a single lineage into two or more distinct and evolutionarily independent lineages—is one of the most important phenomena in biological science. Charles Darwin recognized its preeminence when he chose the title of *The Origin of Species*. But without the underlying knowledge supplied by the modern science of genetics, Darwin was primarily viewing the consequences of speciation, not its underlying causes. Today biologists are actively searching for and finding answers to the many questions about the process of speciation, something biologists have been known to call “the mystery of mysteries.”

IN THIS CHAPTER we will describe what species are and discuss how Earth’s millions of species came into being. We will examine the mechanisms by which a lineage splits into new species and how such separations are maintained. Finally, we will look at different factors that can make speciation a rapid or a very slow process.

CHAPTER OUTLINE

- 23.1 What Are Species?
- 23.2 How Do New Species Arise?
- 23.3 What Happens When Newly Formed Species Come Together?
- 23.4 Why Do Rates of Speciation Vary?

23.1 What Are Species?

Although “species” is a useful and commonly used term in biology, the concept of “species” sometimes varies among different biologists. Biologists are interested in several different aspects of the divergence of biological lineages. Different biologists think about species differently because they ask different questions about species: How can we recognize and identify species? How do new species arise? How do different species remain separate? Why do rates of speciation differ among groups? In answering these questions, biologists focus on different attributes of species, leading to several different ways of thinking about what species are and how they form. Most of the various *species concepts* proposed by biologists are not mutually exclusive—they are just different ways of approaching the question “What are species?”

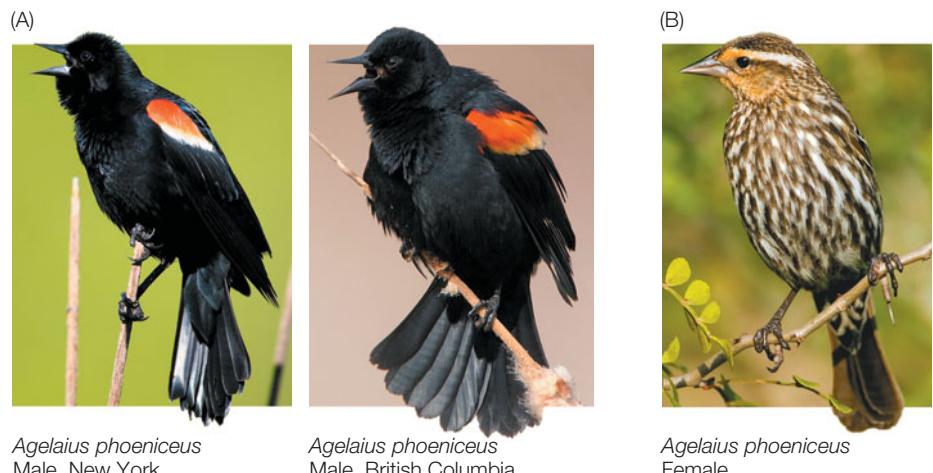
We can recognize many species by their appearance

Biological diversity does not always vary in a smooth, incremental way; groups of organisms often differ in distinct, obvious ways. People have long recognized groups of similar organisms that mate with one another, and there are usually distinct morphological breaks between these groups. Groups of organisms that mate with one another are commonly called *species* (note that this is both the plural and singular form of the word). Someone who is knowledgeable about a group of organisms, such as birds or flowering plants, usually can distinguish the different species found in a particular area simply by looking at them. Standard field guides to birds, mammals, insects, and wildflowers are possible only because many species change little in appearance over large geographic distances. A casual birdwatcher can easily recognize male red-winged blackbirds (*Agelaius phoeniceus*) from the east and west coasts of North America as members of the same species (Figure 23.1A).

More than 250 years ago, the Swedish biologist Carolus Linnaeus developed the binomial system of biological nomenclature by which species are named today (see Section 22.4). Linnaeus described thousands of species, and because he knew nothing about genetics or the mating behavior of the organisms he was naming, he classified them only on the basis of their appearance. Linnaeus differentiated species using a **morphological species concept**, a construct that assumes a species consists of individuals that “look alike,” and that individuals that don’t look alike belong to different species. Although Linnaeus did not know it, members of many of the groups that he classified as species by their appearance look alike because they share many of the alleles that code for their body structures.

23.1 Members of the Same Species Look Alike—or Not

(A) Both of these male red-winged blackbirds are members of the same species—*Agelaius phoeniceus*—even though one is from the eastern United States and the other is from western Canada. Despite the geographic distance, the two individuals are morphologically very similar. (B) Red-winged blackbirds are sexually dimorphic, which means the female of the species looks quite different from the male.



The morphological species concept has limitations, however. In some cases, for instance, not all members of a species look alike. For example, males, females, and young individuals do not always resemble one another closely (**Figure 23.1B**). The morphological species concept is of little use in the case of *cryptic species*, instances in which two or more morphologically indistinguishable species do not interbreed. Biologists therefore cannot always rely on appearance alone in determining whether individual organisms are members of the same or different species. Several additional types of information (especially behavioral and genetic data) are used today to help biologists differentiate species.

Species are reproductively isolated lineages on the tree of life

Evolutionary biologists often think of species as branches on the tree of life, which is known as a **lineage species concept**. Under this concept, each species has a history that starts at a speciation event (one lineage on the tree splits into two) and ends either at extinction or another speciation event, at which time the species produces two daughter species. The process of lineage splitting may be gradual, and thus take thousands of generations to complete. On the other hand, one ancestral lineage may be split into two as a result of a sudden geological event, such as a volcanic eruption that isolates two parts of an ancestral species range. Either way, the lineage species concept treats speciation as the process by which one species splits into two or more daughter species, which thereafter evolve as distinct lineages. The gradual nature of some splitting events means that at a single point in time, the final outcome of the process may not be clear. In these cases it is often impractical to try to decide whether the partially isolated populations of the *incipient species* will continue to diverge and become fully isolated from one another, or perhaps merge again in the future.

The most important component that determines long-term isolation of sexually reproducing lineages from one another is the development of **reproductive isolation**, a state in which two populations can no longer exchange genes. If individuals of a population mate and reproduce with one another, but not with

individuals of other populations, they constitute a distinct group within which genes recombine. In other words, they comprise an independent evolutionary lineage—a separate branch on the tree of life.

It was recognition of the importance of reproductive isolation that brought Ernst Mayr to propose the **biological species concept**: “Species are groups of actually or potentially interbreeding natural populations which are reproductively isolated from other such groups.” The terms “actually” and “potentially” are important elements of the definition. “Actually” says that the individuals live in the same area and interbreed with one another. “Potentially” says that although the individuals do not live in the same area, and therefore cannot interbreed, other information suggests that they would do so if they did get together. This widely used concept of species does not apply to organisms that reproduce asexually, and it is limited to a single point in evolutionary time.

These various concepts of species are not entirely incompatible; they simply emphasize different aspects of species. The morphological species concept emphasizes the practicality of humans recognizing species, although it underestimates or overestimates the actual number of species in some cases. Mayr’s biological species concept emphasizes that reproductive isolation is what keeps sexual lineages on the tree of life separated from one another. The lineage species concept embraces the idea that sexual species are maintained by reproductive isolation, while recognizing the existence of asexual species. The lineage species concept also allows biologists to consider species through evolutionary time, which we will discuss in the next section.

Virtually all species exhibit some degree of genetic recombination among individuals, even if recombination is relatively rare (see Section 21.4). Significant reproductive isolation between species is thus necessary for lineages to remain distinct through evolutionary time. Reproductive isolation is also responsible for the morphological distinctiveness of most species, because genetic mutations that result in morphological changes cannot spread between reproductively isolated species. Therefore, no matter which species concept we emphasize, the evolution of reproductive isolation is important for understanding the origin of species.

23.1 RECAP

Species are distinct lineages on the tree of life. Speciation is usually a gradual process as one lineage divides into two. Over time, lineages of sexual species remain distinct from one another because they have become reproductively isolated.

- Explain how the various species concepts emphasize different attributes of species. **See pp. 482–483**
- Why is the biological species concept not applicable to asexually reproducing organisms?
- Do you understand why reproductive isolation is an important component of all of the species concepts discussed here? **See p. 483**

Although Charles Darwin titled his groundbreaking book *The Origin of Species*, it did not extensively discuss the processes of speciation. He devoted most of his attention to demonstrating that individual species are altered by natural selection over time. We will next discuss the many things that biologists have learned about speciation since Darwin's time.

23.2 How Do New Species Arise?

Not all evolutionary changes result in new species. A single lineage may change over time without giving rise to a new species. Speciation, in contrast, usually requires the evolution of reproductive isolation within a species whose members formerly exchanged genes. But if a genetic change prevents reproduction between individuals of a species, how can such a change spread through a species in the first place?

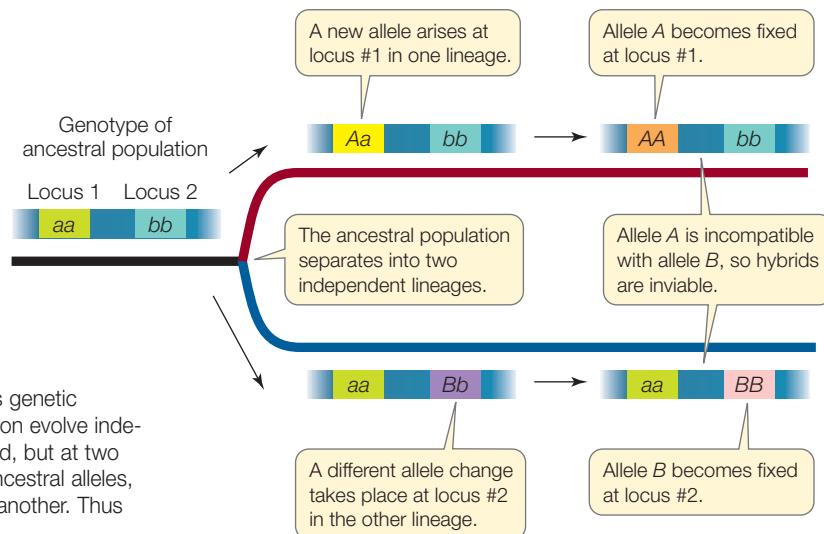
Gene incompatibilities can produce reproductive isolation in two daughter species

If a new allele that causes reproductive incompatibility arises in a population, it cannot spread through the population, because no other individuals are reproductively compatible with the individual that carries the new allele. So how can one reproductively cohesive lineage ever split into two reproductively isolated species? Several early geneticists, including Theodosius Dobzhansky and Hermann Joseph Muller, developed a genetic model to explain this apparent conundrum (**Figure 23.2**).

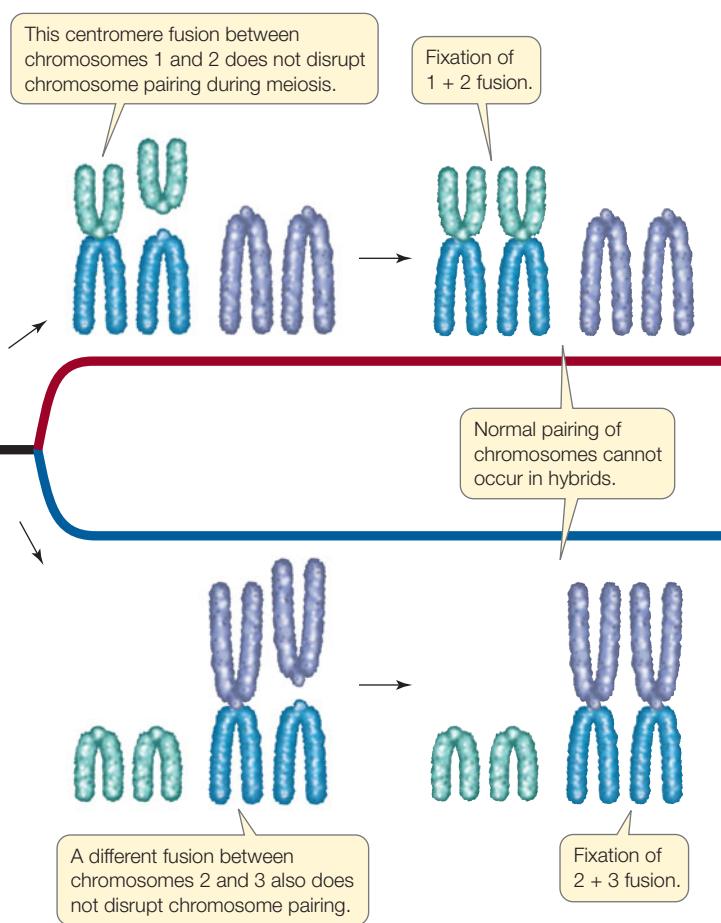
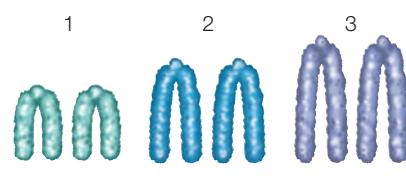
23.2 The Dobzhansky-Muller Model In this simple two-locus genetic model, two populations isolated from the same ancestral population evolve independently. In each descendant lineage, a new allele becomes fixed, but at two different loci. Neither of the new alleles is incompatible with the ancestral alleles, but the two new alleles at the two loci are incompatible with one another. Thus the two descendant lineages are reproductively incompatible.

The Dobzhansky-Muller model is quite simple. First, assume that a single ancestral population is divided into two (by the formation of a new mountain range, for instance). In one of the descendant populations, a new allele (*A*) arises and becomes fixed (see Figure 23.2). In the other population, another new allele (*B*) becomes fixed *at a different gene locus*. Neither new allele at either gene results in any loss of reproductive compatibility. However, the two new forms of these two genes have never occurred together in the same organism. The products of many genes must work together, and it is quite likely that the new protein forms encoded by these two genes are not compatible with one another. If the two populations come back together, they may still be able to interbreed, or **hybridize**. However, the hybrid offspring will have a new combination of genes that may be functionally inferior, or even lethal. Of course, this will not happen with all new combinations of genes, but over time isolated populations will accumulate many allele differences at many gene loci. Thus, we expect that genetic incompatibility will develop through time in the two isolated populations.

Many empirical examples support the Dobzhansky-Muller model, which works not only for pairs of individual genes but also for some kinds of chromosomal rearrangements. For example, bats of the genus *Rhogeessa* exhibit considerable variation in centric fusions of their chromosomes. The chromosomes of the various species contain the same basic chromosomal arms, but in some species two acrocentric (one-armed) chromosomes have fused to form larger, metacentric (two-armed) chromosomes. A polymorphism in this character causes few, if any, problems in meiosis, because the respective chromosomes can still align and assort normally. Therefore, a given centric fusion can become fixed in a population. However, if a *different* centric fusion becomes fixed in a second population, then hybrids between the two populations can no longer produce normal gametes in meiosis (**Figure 23.3**). Most of the closely related species of *Rhogeessa* display different combinations of these centric fusions, and are thereby reproductively isolated from one another.



23.3 Speciation by Centric Fusion In this chromosomal version of the Dobzhansky–Muller model of speciation, two independent centric fusions of one-armed chromosomes occur in two sister lineages. When polymorphic, neither centric fusion event by itself results in major difficulties at meiosis. However, the independent centric fusions are incompatible at meiosis, because the three different chromosomes involved in these fusions cannot pair normally, leading to sterility of the F_1 hybrid. Most of the species in the bat genus *Rhogeessa* differ from one another by such centric fusions.



Reproductive isolation develops with increasing genetic divergence

As pairs of species diverge genetically, they become increasingly reproductively isolated from one another (Figure 23.4). Both the rate at which reproductive isolation develops and the mechanisms that produce it vary from group to group. Nonetheless, reproductive incompatibility has been shown to develop gradually in many groups of plants, animals, and fungi, reflecting the slow pace at which incompatible genes accumulate in each lineage. In some cases, complete reproductive isolation may take millions of years to develop. In other cases (as with the chromosomal fusions of *Rhogeessa* described above), reproductive isolation can develop over just a few generations.

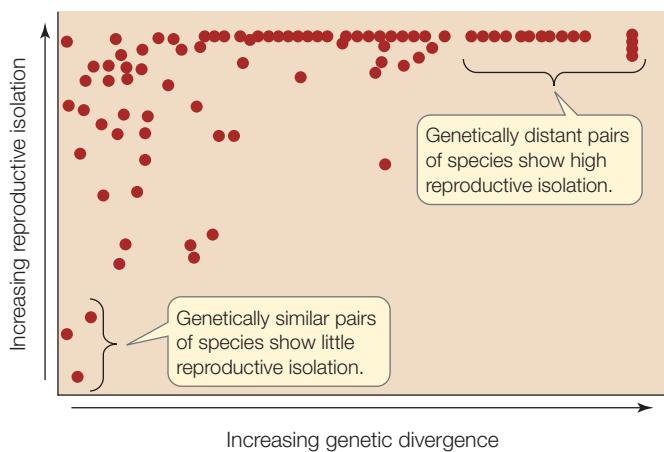
Partial reproductive isolation has evolved in strains of *Phlox drummondii* artificially isolated by humans. In 1835, Thomas Drummond, after whom this species of garden plant is named, collected seeds in Texas and distributed them to nurseries in Europe. Over the next 80 years, the European nurseries established more than 200 true-breeding strains of this phlox, which differed in flower size and color and plant growth form. The plant

breeders did not select directly for reproductive incompatibility between strains, but in subsequent experiments in which fertilization rates were measured and compared, biologists found that reproductive compatibility between strains had been reduced by 14 to 50 percent, depending on the strain.

Many laboratory experiments have demonstrated the gradual evolution of reproductive incompatibility during isolation. We described one such experiment in the opening of this chapter. These laboratory experiments are also consistent with our observations of naturally isolated populations, as described for the unique species of the Cuatro Ciénegas basin in the chapter opening.

Geographic barriers give rise to allopatric speciation

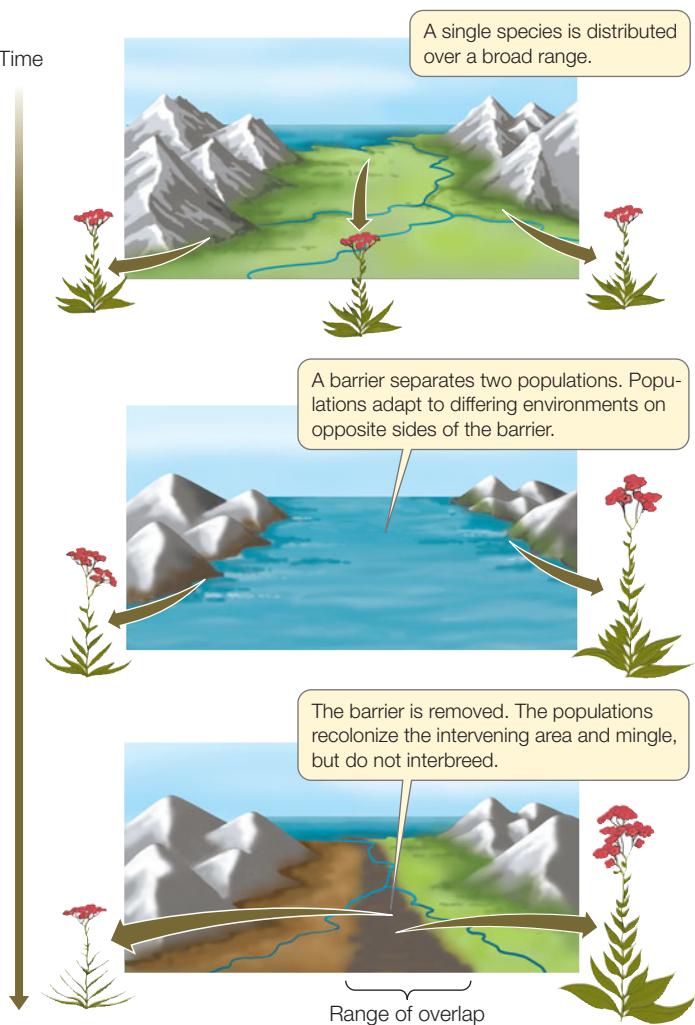
Speciation that results when a population is divided by a physical barrier is known as **allopatric speciation** (Greek *allo*, “other”; *patra*, “homeland”), also called *geographic speciation* (Figure 23.5). Allopatric speciation is thought to be the dominant mode of speciation in most groups of organisms. The physical barrier that divides the range of a species may be a water body, a mountain range, or other inhospitable habitat for terrestrial organisms, or dry land for aquatic organisms. Barriers can form when continents drift, sea levels rise, glaciers advance and retreat, or climates change. These processes continue to generate



23.4 Reproductive Isolation Increases with Time of Separation

This positive relationship between genetic distance and reproductive isolation, shown here for pairs of *Drosophila* species, has been observed in many groups of plants, animals, and fungi (each dot represents a comparison of one species pair).

Time



23.5 Allopatric Speciation Allopatric speciation may result when a population is divided into two separate populations by a physical barrier, such as rising sea levels.

physical barriers today. The populations separated by such barriers are often, but not always, initially large. They evolve differences for a variety of reasons, including genetic drift (see Section 21.1), but especially because the environments in which they live are, or become, different.

Allopatric speciation may also result when some members of a population cross an existing barrier and found a new, isolated population. Many of the more than 800 species of *Drosophila* found in the Hawaiian archipelago are restricted to a single island. We know that these species are the descendants of new populations founded by individuals dispersing among the islands because the closest relative of a species on one island is often a species on a neighboring island rather than a species on the same island. Biologists who have studied the chromosomes of these fruit flies estimate that speciation in this group of *Drosophila* has resulted from at least 45 such founder events (Figure 23.6).

The 14 species of finches found in the Galápagos, an archipelago 1,000 kilometers off the coast of Ecuador, were gener-

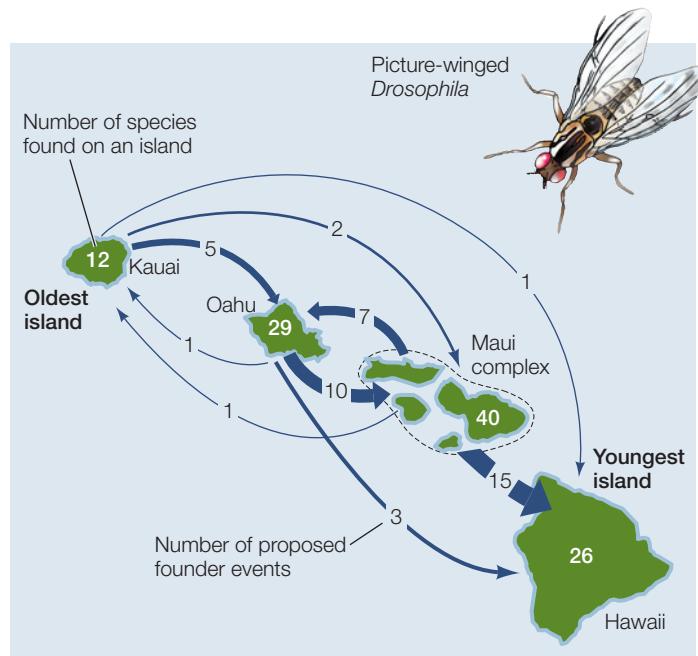
ated by allopatric speciation. Darwin's finches (as they are usually called, because Darwin was the first scientist to study them) arose in the Galápagos from a single South American species that colonized the islands. Today the 14 species differ strikingly from their closest mainland relative and from one another (Figure 23.7). The islands of the Galápagos archipelago are sufficiently far apart that finches move among them only infrequently. In addition, environmental conditions differ from island to island. Some are relatively flat and arid; others have forested mountain slopes. Finch populations on different islands have differentiated over millions of years to the point that, when occasional immigrants arrive from other islands, they either do not breed with the residents or, if they do, the resulting offspring do not survive as well as the offspring of established island residents. The genetic distinctness and cohesiveness of the different finch species are thus maintained.

yourBioPortal.com

GO TO Animated Tutorial 23.1 • Founder Events and Allopatric Speciation

Sympatric speciation occurs without physical barriers

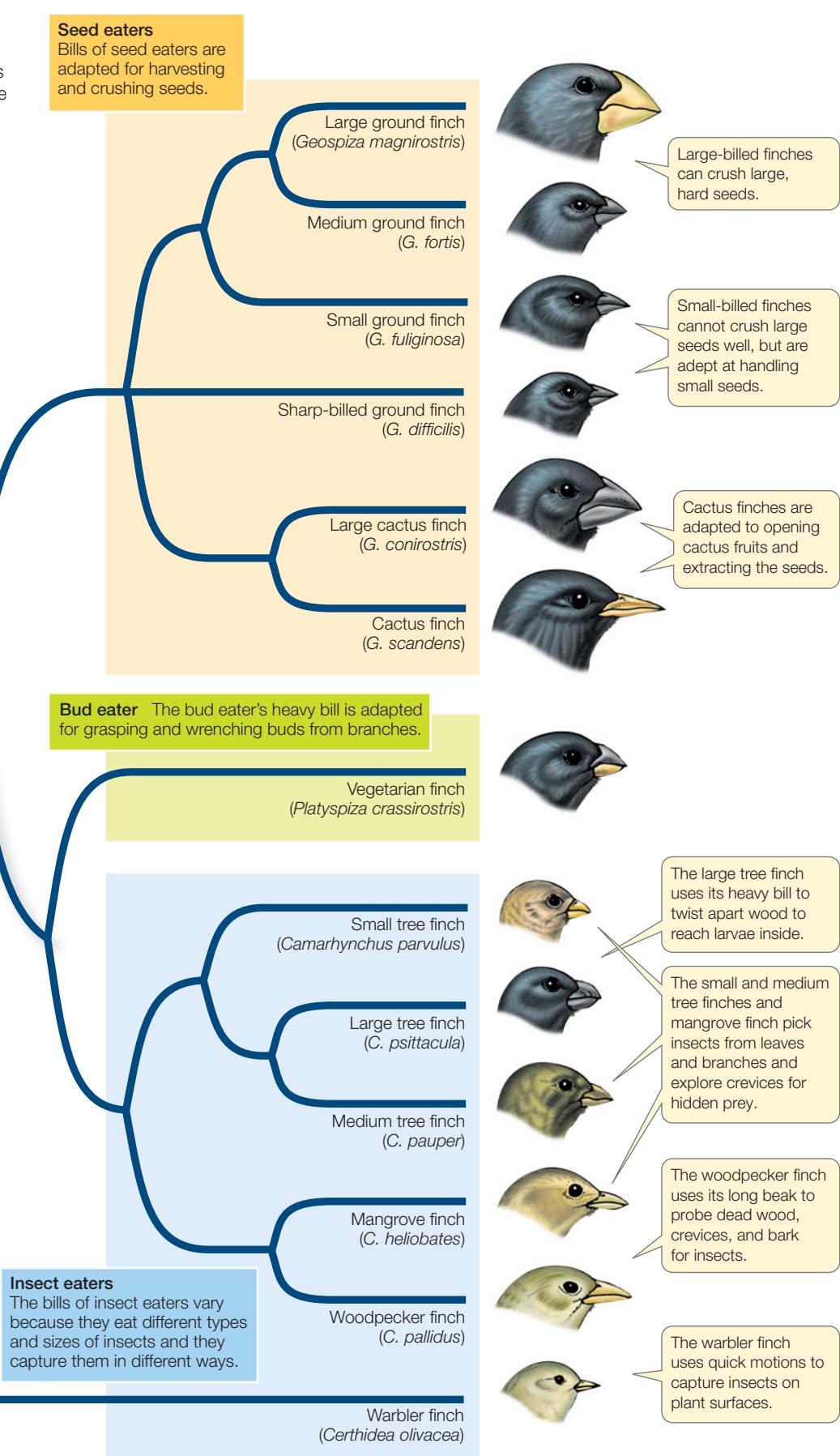
Although physical isolation is usually required for speciation, under some circumstances speciation can occur without it. A partition of a gene pool without physical isolation is called **sympatric speciation** (Greek *sym*, "together with"). Given that speciation is usually a gradual process, how can reproductive isolation develop when individuals have frequent opportunities



23.6 Founder Events Lead to Allopatric Speciation The large number of species of picture-winged *Drosophila* in the Hawaiian Islands is the result of founder events: the founding of new populations by individuals dispersing among the islands. The islands, which were formed in sequence as Earth's crust moved over a volcanic "hot spot," vary in age.

23.7 Allopatric Speciation among Darwin's Finches

Darwin's Finches The descendants of the ancestral finch that colonized the Galápagos archipelago several million years ago evolved into 14 different species whose members are variously adapted to feed on seeds, buds, and insects. (The fourteenth species, not pictured here, lives in Cocos Island, farther north in the Pacific Ocean.)



to mate with one another? Sympatric speciation may occur with some form of disruptive selection in which certain genotypes have a preference for distinct microhabitats where mating takes place. The experiment described in the opening of this chapter shows that this kind of disruptive selection can take place in the laboratory, but does it also occur in nature?

Sympatric speciation via disruptive selection appears to be happening in the apple maggot fly (*Rhagoletis pomonella*) in eastern North America. Until the mid-1800s, *Rhagoletis* flies courted, mated, and deposited their eggs only on hawthorn fruits. About 150 years ago, some *Rhagoletis* flies began to lay their eggs on apples, which European immigrants had introduced into eastern North America. Apple trees are closely related to hawthorns, but the smell of the fruits differs, and the apple fruits appear earlier than those of hawthorns. Some early-emerging female *Rhagoletis* laid their eggs on apples and evolved a genetic preference for the smell of apples. Their offspring inherited this genetic preference for apples for mating and egg deposition. When the offspring sought out apple trees for these purposes, they mated with other flies reared on apples, which shared the same preferences.

Today the two groups of *Rhagoletis pomonella* in the eastern United States may be on the way to becoming distinct species. One group mates and lays eggs primarily on hawthorn fruits, the other on apples. The two incipient species are partly reproductively isolated because they mate primarily with individuals raised on the same fruit and because they emerge from their pupae at different times of the year. In addition, the apple-feeding flies have evolved so that they now grow more rapidly on apples than they originally did.

Sympatric speciation via ecological isolation, as appears to be happening in *Rhagoletis pomonella*, may be widespread among insects, many of which feed on only a single plant species. The most common means of sympatric speciation, however, is **polyploidy**, or the duplication of sets of chromosomes within individuals (see Section 11.5). Polyploidy can arise either from chromosome duplication in a single species (**autopolyploidy**) or from the combining of the chromosomes of two different species (**allopolyploidy**).

An autopolyploid individual originates when (for example) two accidentally unreduced diploid gametes (with

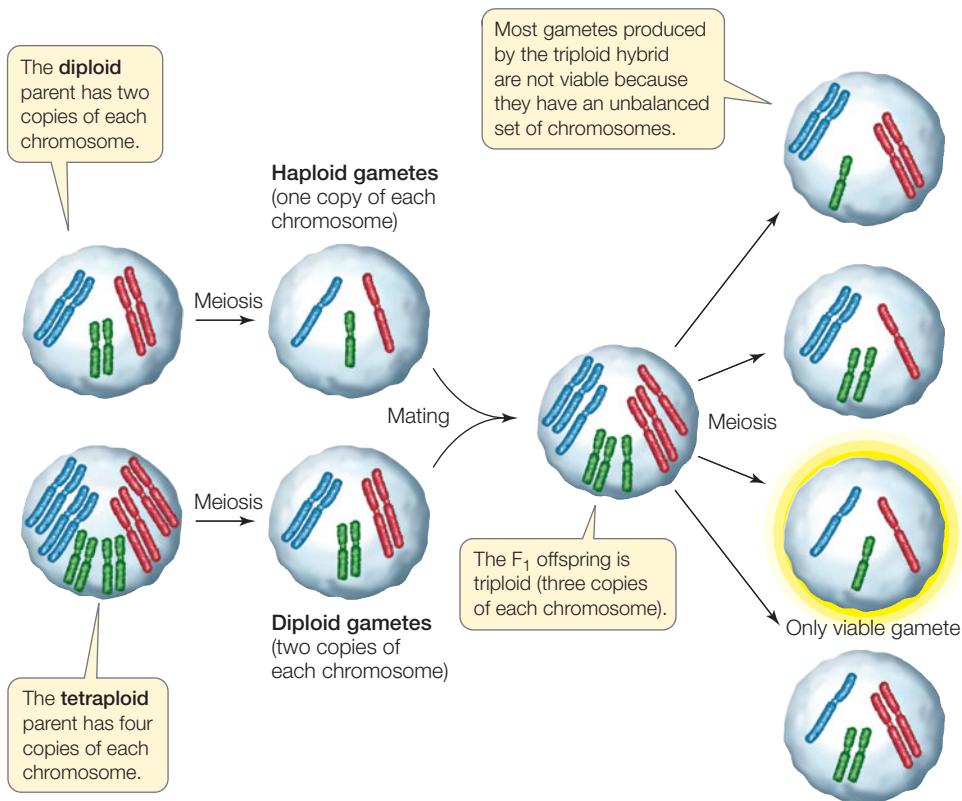
23.8 Tetraploids Are Repproductively Isolated from Their Diploid Ancestors

Even if the triploid offspring of a diploid and a tetraploid parent survives and reaches sexual maturity, most of the gametes it produces have aneuploid (unbalanced) numbers of chromosomes. Such triploid individuals are effectively sterile. (For simplicity, the diagram shows only three chromosomes; most species have many more than that.)

two sets of chromosomes) combine to form a tetraploid individual (with four sets of chromosomes). Tetraploid and diploid individuals of the same species are reproductively isolated because their hybrid offspring are triploid and are usually sterile; they cannot produce viable gametes because their chromosomes do not segregate evenly during meiosis (Figure 23.8). So a tetraploid individual cannot produce viable offspring by mating with a diploid individual—but it *can* do so if it self-fertilizes or mates with another tetraploid. Thus polyploidy can result in complete reproductive isolation in two generations—an important exception to the general rule that speciation is a gradual process.

Allopolyploids may also be produced when individuals of two different (but closely related) species interbreed. Such hybridization often disrupts normal meiosis, which can result in chromosomal doubling. Allopolyploids are often fertile because each of the chromosomes has a nearly identical partner with which to pair during meiosis.

Speciation by polyploidy has been particularly important in the evolution of plants. Botanists estimate that about 70 percent of flowering plant species and 95 percent of fern species are the result of recent polyploidization. Some of these arose from hybridization between two species, followed by chromosomal duplication and self-fertilization. Many other species diverged from polyploid ancestors, so the new species also shared the duplicated sets of chromosomes. New species may arise by means of polyploidy more easily among plants than among animals because plants of many species can reproduce by self-fertilization. In addition, if polyploidy arises in several offspring of a single parent, the siblings can fertilize one another.



23.2 RECAP

Allopatric speciation results from the separation of populations by geographic barriers; it is the dominant mode of speciation among most groups of organisms. Sympatric speciation may result from ecological isolation, but among plants and some animals, polyploidy is the most common cause of sympatric speciation.

- How can speciation via polyploidy happen in two generations? See p. 488
- Explain why an effective barrier to gene flow for one species may not effectively isolate another species.

yourBioPortal.com

GO TO Animated Tutorial 23.2 • Speciation Mechanisms

Polyploidy, as we have just seen, can result in a new species that is completely reproductively isolated from its parent species in two generations, but most populations separated by a physical barrier become reproductively isolated only very slowly. Let's see how reproductive isolation may become established once two populations have separated from each other.

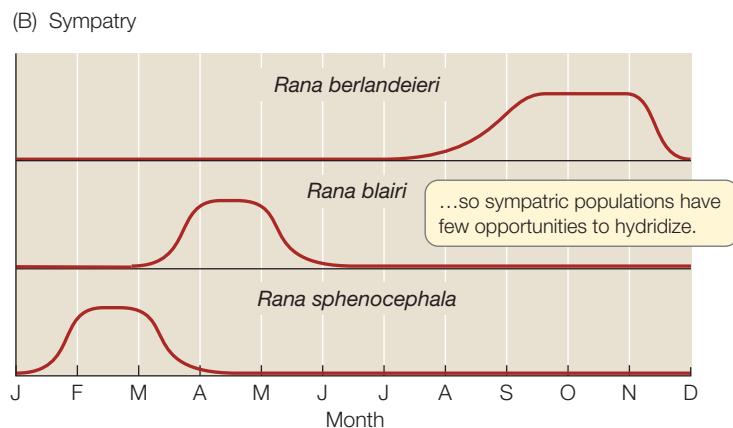
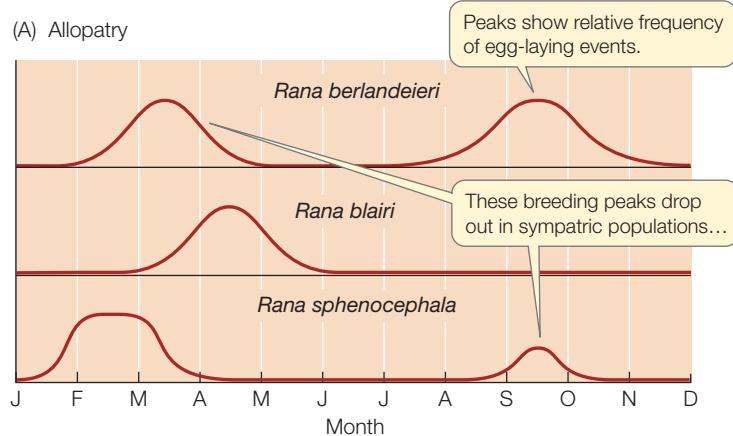
23.3 What Happens When Newly Formed Species Come Together?

As discussed in the previous section, once a barrier to gene flow is established, reproductive isolation can develop through genetic divergence. Over many generations, genetic differences accumulate that reduce the probability that members of the two populations can mate and produce viable offspring. In this way, reproductive isolation evolves as a by-product of the genetic changes in the two populations. What types of mechanisms prevent or reduce gene flow between populations, leading to reproductive isolation? Reproductive isolating mechanisms fall into two major categories: **prezygotic reproductive barriers** act before fertilization to prevent individuals of different species or populations from mating, whereas **postzygotic reproductive barriers** act after fertilization to prevent the development of viable offspring, or to reduce the offspring's fertility.

Prezygotic barriers prevent fertilization

Prezygotic mechanisms come into play before fertilization and can involve several kinds of reproductive isolation.

HABITAT ISOLATION When individuals of different species evolve genetic preferences for different habitats in which they live or mate, they may never come into contact during their respective mating periods. The *Rhagoletis* flies in the eastern United States (discussed in Section 23.2) experienced such habitat isolation, as did the *Drosophila* in the experiment described in the opening of this chapter.



23.9 Temporal Isolation in the Breeding Seasons of Three Species of Frogs (A) The peak breeding seasons of three species of *Rana* overlap when the species are physically separated (allopatry). (B) When two or more species of *Rana* occupy the same territory (sympatry), overlap between peak breeding seasons of each species is greatly reduced or eliminated. In areas where only one species is found, the breeding seasons are broader. Selection against hybridization in areas of overlap helps reinforce the prezygotic isolating mechanism.

TEMPORAL ISOLATION Many organisms have distinct mating seasons. If two closely related species breed at different times of the year (or different times of day), the two may never have an opportunity to hybridize. For example, in sympatric populations of three closely related leopard frogs, each species breeds at a different time of year (Figure 23.9). Although there is some overlap in the breeding seasons, the opportunities for hybridization are minimized.

MECHANICAL ISOLATION Differences in the sizes and shapes of reproductive organs may prevent the union of gametes from different species. With animals, this may involve a match in the shape of reproductive organs between males and females, so that reproduction between species with mismatching structures is not physically possible. In plants, the mechanical isolation between species may involve a pollinator. For example, orchids of the genus *Cryptostylis* produce flowers that look and smell like the females of particular species of wasps (Figure 23.10). When a male wasp visits and attempts to mate with the flower (thinking it is a female wasp), the mating action results in transfer of

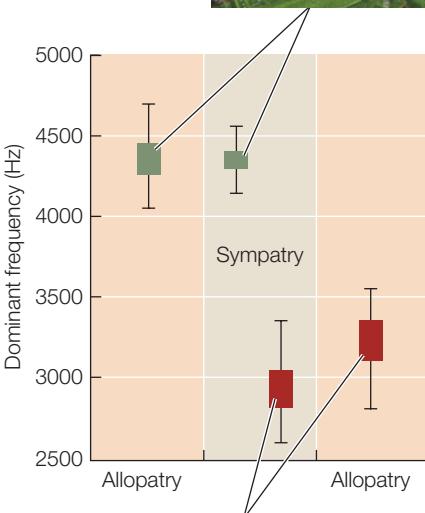
23.10 Mechanical Isolation through Mimicry Many orchid species maintain reproductive isolation because their flowers look and smell like a specific species of bee or wasp, inducing copulatory actions on the part of that specific pollinator insect. The placement of the anthers and stigmas on the flower results in transfer of pollen from the flower to the insect that “mates” with it, and then from the insect to the next flower with which it attempts to mate. Shown here are an Australian orchid (*Cryptostylis* sp.) and its pollinator, a male wasp of the genus *Lissopimpla*.

pollen between the flower and wasp as a result of appropriately configured anthers and stigmas on the flower. Insects that visit the flower but do not attempt to mate with it do not trigger the transfer of pollen between the insect and flower.

BEHAVIORAL ISOLATION Individuals of a species may reject, or fail to recognize, individuals of other species as potential mating partners. For example, the breeding calls of male frogs quickly diverge between related species (Figure 23.11). Female frogs respond to and approach calls from males of their own species, but ignore the calls of even closely related species.



Gastrophryne olivacea



Gastrophryne carolinensis

23.11 Behavioral Isolation in the Mating Calls of Male Frogs The males of most species of frogs produce species-specific calls. The calls of the two closely related frog species in this figure differ in their dominant frequency (a high-frequency sound wave results in a high-pitched sound; a low frequency results in a low-pitched sound). Female frogs are attracted to the calls of males of their own species. Note that the calls of the two species are more distinct in areas of sympatry than in areas of allopatry (an example of reinforcement).

Sometimes the mate choice of one species is mediated by the behavior of individuals of other species. For example, whether two plant species hybridize may depend on the food preferences of their pollinators. The floral traits of plants, including their color and shape, can enhance reproductive isolation either by influencing which pollinators are attracted to the flowers or by altering where pollen is deposited on the bodies of pollinators. A plant whose flowers are pendant (Figure 23.12A) will be pollinated by an animal with different physical characteristics than a plant in which the flowers grow upright (Figure 23.12B). Because each pollinator prefers (and is adapted to) a different type of flower, the pollinators rarely transfer pollen from one plant species to the other.

Such isolation by pollinator behavior is seen in the case of two sympatric species of columbines (*Aquilegia*) in the mountains of California that have diverged in flower color, structure, and orientation. *Aquilegia formosa* has pendant flowers with short spurs (spikelike, nectar-containing structures) and is pollinated by hummingbirds (Figure 23.12C). *A. pubescens* has upright, lighter-colored flowers with long spurs and is pollinated by hawkmoths (Figure 23.12D). The difference in pollinators means that these two species are effectively reproductively isolated even though they populate the same geographic range.

GAMETIC ISOLATION Sperm of one species may not attach to the eggs of another species because the eggs do not release the appropriate attractive chemicals, or the sperm may be unable to penetrate the egg because the two gametes are chemically in-

(A)



(B)

(C) *Aquilegia formosa*(D) *Aquilegia pubescens*

compatible. Thus, even though individuals of the two species may attempt to mate, the gametes never fuse into a zygote. For example, gametic isolation has arisen between species of sea urchins. A protein known as bindin occurs in sea urchin sperm and functions in attaching (“binding”) the sperm to eggs. All sea urchins produce this egg-recognition protein, but the gene sequence diverges rapidly between species. The sperm protein evolves so that it will only bind to eggs of the same species, thus preventing interspecific hybridization.

Postzygotic barriers can isolate species after fertilization

If individuals of two different populations lack complete prezygotic reproductive barriers, postzygotic reproductive barriers may prevent the species from merging. Genetic differences that accumulate while the populations are isolated from each other may reduce the survival and reproduction of hybrid offspring in any of several ways:

- *Low hybrid zygote viability.* Hybrid zygotes may fail to mature normally, either dying during development or developing such severe abnormalities that they cannot mate as adults.
- *Low hybrid adult viability.* Hybrid offspring may simply have lower survivorship than offspring resulting from within-population matings.

23.12 Floral Morphology is Associated with Pollinator Morphology

(A) This hummingbird's morphology and behavior are adapted to approach plants whose flowers are pendant (hanging downward). (B) The nectar-extracting proboscis of this hawkmoth is adapted to flowers that grow upright. (C) *Aquilegia formosa* flowers are normally pendant and are pollinated by hummingbirds. (D) Flowers of *A. pubescens* are normally upright, which facilitates pollination by hawkmoths. In addition, the long floral spurs appear to restrict access by some other potential pollinators.

- *Hybrid infertility.* Hybrids may mature normally but be infertile. For example, the offspring of matings between horses and donkeys—mules—are healthy but sterile; they produce no descendants.

Although natural selection does not directly favor the evolution of postzygotic reproductive barriers, if hybrid offspring survive poorly, natural selection may favor the evolution of prezygotic barriers. This happens because individuals that mate with individuals of the other population will leave fewer surviving descendants than individuals that mate only within their own population. In this case, individuals that can avoid mating with members of the other population have a selective advantage, and any trait that favors such avoidance will be favored by natural selection. Such strengthening of prezygotic barriers is known as **reinforcement**.

Donald Levin of the University of Texas noticed that individuals of *Phlox drummondii* in most of the range of the species in Texas have pink flowers. Where *P. drummondii* is sympatric with the pink-flowered *P. cuspidata*, however, *P. drummondii* has red flowers. No other *Phlox* species has red flowers. Levin performed an experiment whose results showed that reinforcement might explain the evolution of red flowers where the two species are sympatric (**Figure 23.13**).

Reinforcement can also be detected by comparing sympatric and allopatric populations of potentially hybridizing species. If reinforcement is occurring, then sympatric pairs of closely related species should evolve more effective prezygotic reproductive barriers than do allopatric populations of the same species. The examples of temporal isolation shown in Figure 23.9 and of behavioral isolation shown in Figure 23.11 illustrate reinforcement of prezygotic barriers. The breeding seasons of the sympatric populations of frogs (Figure 23.9) overlap much less than do those of the corresponding allopatric populations. Similarly, the frequencies of the frog mating calls illustrated in Figure 23.11 are more divergent in sympatric populations than in allopatric populations. In both cases, there appears to have been selection against hybrids in areas of sympatry, so individuals that do not produce hybrids are more likely to leave more genes to future generations.

INVESTIGATING LIFE

23.13 Flower Color Reinforces a Reproductive Barrier in *Phlox*

Most *Phlox drummondii* flowers are pink, but in regions where they are sympatric with *P. cuspidata*—which is always pink—most *P. drummondii* individuals are red. Most pollinators preferentially visit flowers of one color or the other. In this experiment, Donald Levin explored whether flower color reinforces a prezygotic reproductive barrier, lessening the chances of interspecific hybridization.

HYPOTHESIS

Red-flowered *P. drummondii* are less likely to hybridize with *P. cuspidata* than are pink-flowered *P. drummondii*.

METHOD

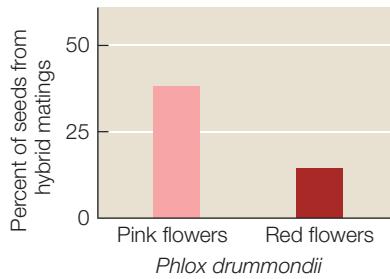
1. Introduce equal numbers of red- and pink-flowered *P. drummondii* individuals into an area with many pink-flowered *P. cuspidata*.



2. After the flowering season ends, measure hybridization by assessing the genetic composition of the seeds produced by *P. drummondii* plants of both colors.

RESULTS

Of the seeds produced by pink-flowered *P. drummondii*, 38% were hybrids with *P. cuspidata*. Only 13% of the seeds produced by red-flowered individuals were genetic hybrids.



CONCLUSION

P. drummondii and *P. cuspidata* are less likely to hybridize if the flowers of the two species differ in color.

FURTHER INVESTIGATION:

This experiment did not address the probable reproduction advantages for individual *Phlox* plants of donating and receiving primarily intraspecific pollen. Can you design an experiment to measure such an advantage?

Go to [yourBioPortal.com](#) for original citations, discussions, and relevant links for all INVESTIGATING LIFE figures.

During the process of reinforcement, closely related species may form hybrids in areas where their ranges overlap, and they may continue to do so for many years. Let's examine what happens when reproductive barriers do not completely prevent individuals from different populations from mating and producing offspring.

Hybrid zones may form if reproductive isolation is incomplete

If contact is reestablished between formerly isolated populations before complete reproductive isolation has developed, members of the two populations may interbreed. Three outcomes of such interbreeding are possible:

- If hybrid offspring are as fit as those resulting from matings within each population, hybrids will mate with individuals of both parental species. The gene pools will gradually become completely mixed, resulting in one species.
- If hybrid offspring are less fit, complete reproductive isolation may evolve as reinforcement strengthens prezygotic reproductive barriers.
- Even if hybrid offspring are at some disadvantage, a narrow **hybrid zone**—a region in which genetically distinct populations come together and produce offspring of mixed ancestry—may develop in the absence of reinforcement, or before reinforcement is complete.

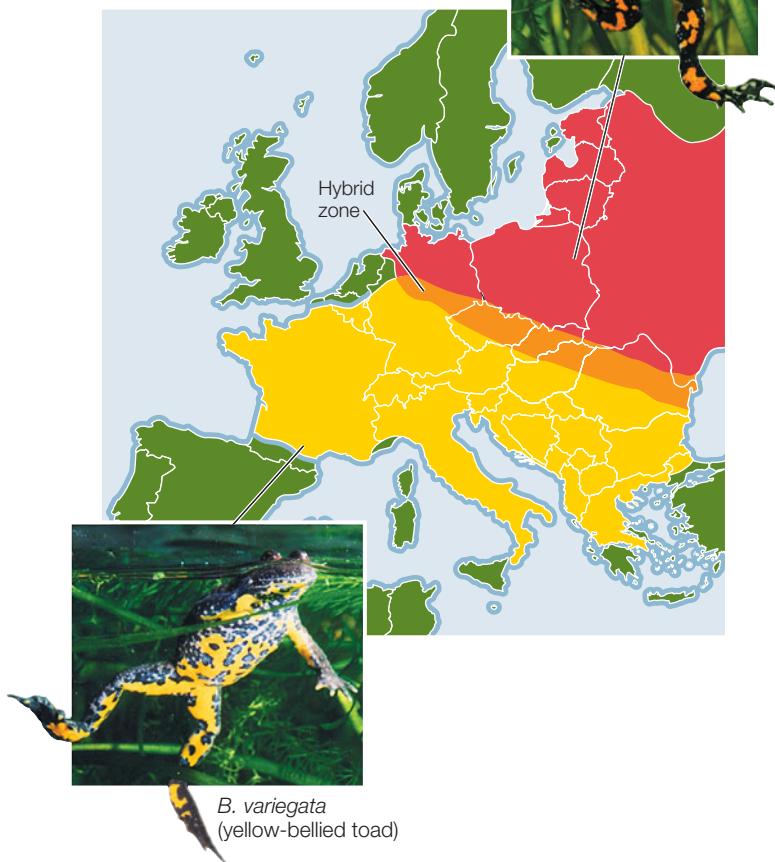
When a hybrid zone first forms, most hybrids are offspring of crosses between purebred individuals of the two populations. However, subsequent generations include a variety of individuals with different proportions of their genes derived from the original two populations. Thus hybrid zones contain recombinant individuals resulting from many generations of hybridization. Detailed genetic studies can tell us much about why hybrid zones may be narrow and stable for long periods of time.

The hybrid zone between two species of European toads of the genus *Bombina* has been studied intensively. The fire-bellied toad (*B. bombina*) lives in eastern Europe; the closely related yellow-bellied toad (*B. variegata*) lives in western and southern Europe. The ranges of the two species overlap in a long but very narrow zone stretching 4,800 kilometers from eastern Germany to the Black Sea (Figure 23.14). Hybrids between the two species suffer from a range of defects, many of which are lethal. Those hybrids that survive often have skeletal abnormalities, such as misshapen mouths, ribs that are fused to vertebrae, and a reduced number of vertebrae. By following the fates of thousands of toads from the hybrid zone, investigators found that a hybrid toad is on average only half as fit as a purebred individual. The

23.14 Hybrid Zones May Be Long and Narrow

Narrow The narrow zone in Europe where fire-bellied toads meet and hybridize with yellow-bellied toads stretches across Europe. This hybrid zone has been stable for hundreds of years and has never expanded, and no reinforcement has evolved.

B. bombina (fire-bellied toad)



hybrid zone thus remains narrow because there is strong selection against hybrids, and because adult toads do not move over long distances. The zone has persisted for hundreds of years, however, because many individuals of both species continue to move short distances into it, constantly replenishing the hybrid population.

23.3 RECAP

Reproductive isolation may result from prezygotic or postzygotic reproductive barriers. Lower fitness of hybrids in contact zones can lead to the reinforcement of prezygotic reproductive barriers.

- Describe various kinds of prezygotic and postzygotic reproductive barriers. **See pp. 489–491**
- Why is reinforcement of prezygotic barriers likely if hybrid offspring survive more poorly than offspring produced by within-population matings? **See p. 491**

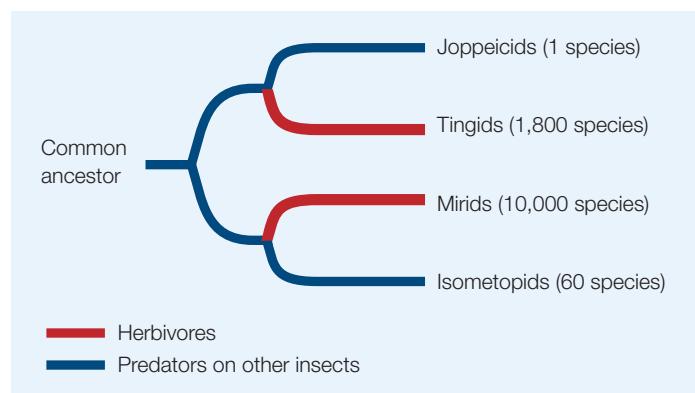
Some groups of organisms have many species, others only a few. Hundreds of species of *Drosophila* evolved in the small area of the Hawaiian Islands over about 20 million years. In contrast, there are only a few species of horseshoe crabs in the world, and only one species of ginkgo tree, even though these latter groups have persisted for hundreds of millions of years. Why do different groups of organisms have such different rates of speciation?

23.4 Why Do Rates of Speciation Vary?

Rates of speciation (the proportion of existing species that split to form new species over a given period of time) vary greatly because many factors influence the likelihood that a lineage will split to form two or more species. What are some of the factors that influence the probability of a given lineage splitting into two?

Populations of species that have specialized diets may be more likely to diverge than are populations that have generalized diets. To investigate the effects of diet on rates of speciation, Charles Mitter and colleagues compared species richness in some closely related groups of true bugs (hemipterans). The common ancestor of these groups was a predator that fed on other insects, but a dietary shift to herbivory (eating plants) evolved at least twice in the groups under study. The herbivorous groups have many more species than those that are predatory (Figure 23.15). Herbivorous bugs typically specialize on one or a few closely related species of plants, whereas predatory bugs tend to feed on many different species of insects. High diversity of host plants can thus lead to a correspondingly high diversity in the herbivorous specialists.

Speciation rates in plants are faster in animal-pollinated than in wind-pollinated plants. Animal-pollinated groups have, on average, 2.4 times as many species as related groups pollinated by wind. Among animal-pollinated plants, speciation rates are correlated with pollinator specialization. In columbines (*Aquilegia*), the rate of evolution of new species has been about three times faster in lineages that have long nectar spurs than in lineages that lack spurs. Why do nectar spurs increase the spe-



23.15 Dietary Shifts Can Promote Speciation Herbivorous groups of hemipteran insects have speciated several times faster than closely related predatory groups.

ciation rate? Apparently it is because having longer spurs restricts the number of pollinator species that visit the flowers, thus increasing opportunities for reproductive isolation (see Figure 23.12).

The mechanisms of sexual selection (see Section 21.2) appear to result in increased rates of speciation. Some of the most striking examples of sexual selection are found in birds with promiscuous mating systems. Bird-watchers travel thousands of miles to Papua New Guinea to witness the mating displays of male birds of paradise, some of which have long, brightly colored tail feathers and look distinctly different than the females (*sexual dimorphism*). In many of these 33 species, males assemble at display grounds, called *leks*, and females come there to choose a male with whom to copulate. After mating, the females leave the display grounds, build their nests, lay their eggs, and feed their offspring with no help from the males. The males remain to court more females (Figure 23.16A).

The closest relatives of the birds of paradise are the manucodes. Male and female manucodes differ only slightly in size and plumage (so they are *sexually monomorphic*). They form monogamous pair bonds, and both sexes contribute to raising the young. There are only 5 species of manucodes (Figure 23.16B), compared with 33 species of birds of paradise. This

(A) *Paradisaea minor*



(B) *Manucodia comrii*



23.16 Sexual Selection in Birds Can Lead to Higher Speciation Rates

Rates (A) Birds of paradise and (B) manucodes are closely related bird groups of the South Pacific. However, speciation rates are much higher among the sexually dimorphic, polygynous birds of paradise (33 species) than among manucodes (5 species).

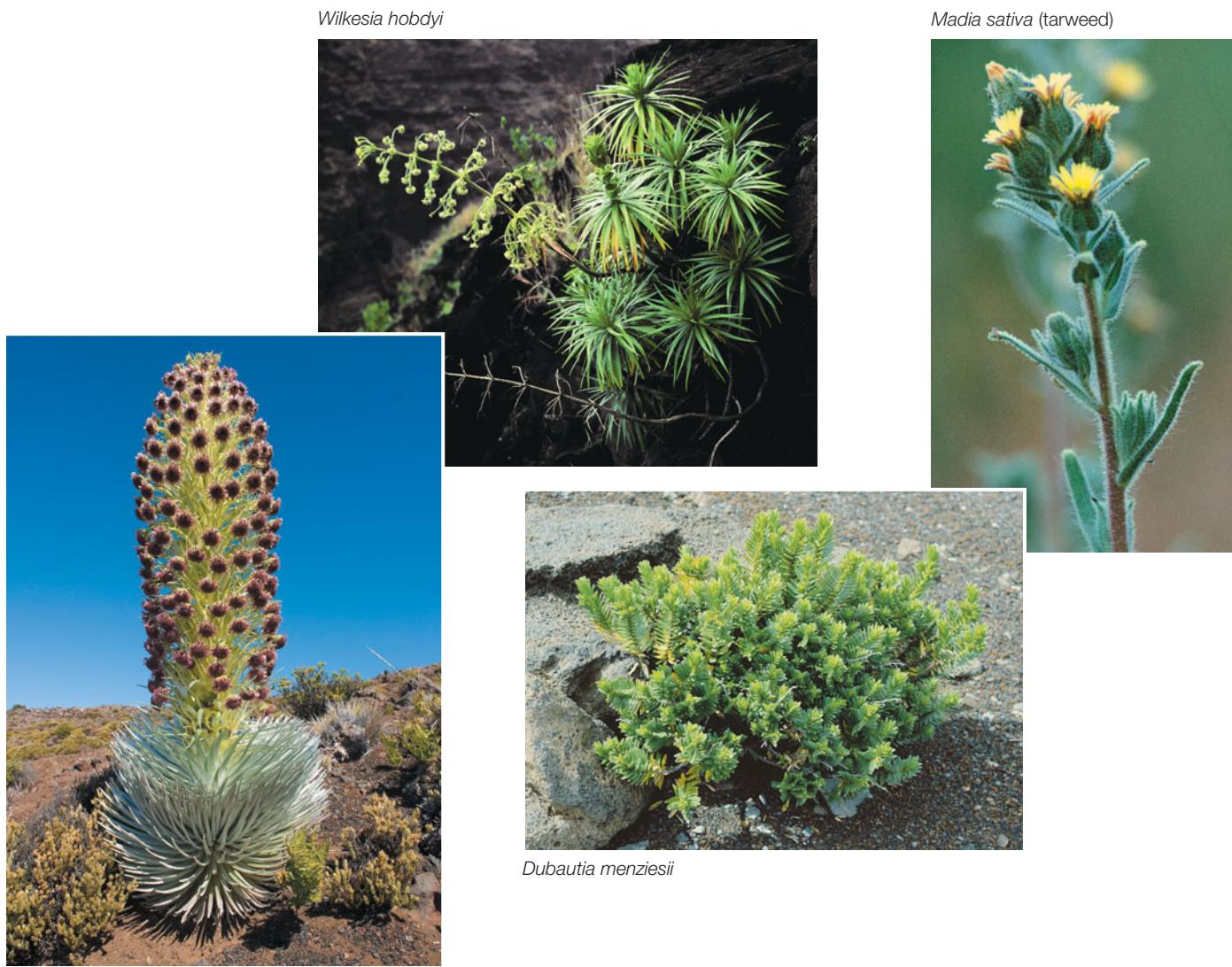
is just one comparison, and by itself would not be convincing evidence that sexually dimorphic clades of birds have higher rates of speciation than do monomorphic clades. However, when biologists compare all the examples of birds in which one clade is sexually dimorphic, and the most closely related clade is sexually monomorphic, the sexually dimorphic clades are significantly more likely to contain more species. But why would sexual dimorphism be associated with a higher rate of speciation?

Animals with complex sexually selected behaviors are likely to form new species at a high rate because they make sophisticated discriminations among potential mating partners. They distinguish members of their own species from members of other species, and they make subtle discriminations among members of their own species on the basis of size, shape, appearance, and behavior. Such discriminations can greatly influence which individuals are most successful in mating and producing offspring, and may lead to rapid evolution of prezygotic reproductive barriers among populations.

Speciation rates are usually higher in species with poor dispersal abilities than in species with good dispersal abilities, because even narrow barriers can be effective in dividing a species whose members are highly sedentary. The Hawaiian Islands have about 1,000 species of land snails, many of which are restricted to a single valley. Because snails move only short distances, the high ridges that separate the valleys are effective barriers to their dispersal.

The proliferation of a large number of daughter species from a single ancestor is called an **evolutionary radiation**. If the rapid proliferation of species results in an array of species that live in a variety of environments and differ in the characteristics they use to exploit those environments, the radiation is said to be **adaptive**. Several remarkable adaptive radiations have occurred in the Hawaiian Islands. In addition to its 1,000 species of land snails, the native biota of the Hawaiian Islands includes 1,000 species of flowering plants, 10,000 species of insects, and more than 100 bird species. However, there were no amphibians, no terrestrial reptiles, and only one native terrestrial mammal—a bat—on the islands until humans introduced additional species. The 10,000 known native species of insects on Hawaii are believed to have evolved from only about 400 immigrant species; only 7 immigrant species are believed to account for all the native Hawaiian land birds. Similarly, as we saw earlier in this chapter, an adaptive radiation in the Galápagos archipelago resulted in the 14 species of Darwin’s finches, which differ strikingly in the size and shape of their bills and, accordingly, in the food resources they use (see Figure 23.7).

The 28 species of Hawaiian sunflowers called silverswords are an impressive example of an adaptive radiation in plants. DNA sequences show that these species share a relatively recent common ancestor with a species of tarweed from the Pacific coast of North America (Figure 23.17). Whereas all mainland tarweeds are small, upright herbs (nonwoody plants), the silverswords include prostrate and upright herbs, shrubs, trees, and vines. Silversword species occupy nearly all the habitats of the Hawaiian Islands, from sea level to above timberline

*Argyroxiphium sandwicense*

23.17 Rapid Evolution among Hawaiian Silverswords The Hawaiian silverswords, three closely related genera of the sunflower family, are believed to have descended from a single common ancestor (a plant similar to the tarweed; upper right) that colonized Hawaii from the Pacific coast of North America. The four plants shown here are more closely related than they appear to be based on their morphology.

in the mountains. Despite their extraordinary morphological diversification, the silverswords are genetically very similar.

The Hawaiian silverswords are more diverse in size and shape than the mainland tarweeds because the tarweed ancestors first arrived on islands that harbored very few plant species. In particular, there were few trees and shrubs, because such large-seeded plants rarely disperse to oceanic islands. Trees and shrubs have evolved from nonwoody ancestors on many oceanic islands. On the mainland, however, tarweeds live in ecological communities that contain many tree and shrub species in lineages with long evolutionary histories. In those environments, opportunities to exploit the “tree” way of life had already been preempted.

23.4 RECAP

Dispersal ability, dietary specialization, and mechanisms of sexual selection affect rates of speciation. Speciation rates in plants can depend on mechanisms of pollination. Open ecological niches present opportunities for evolutionary radiations.

- Explain how pollinator specialization in plants and sexual selection in animals can increase rates of speciation. **See pp. 493–494**
- Why do adaptive radiations often occur when a founder species invades an isolated geographic area? **See p. 494**

The processes described in this chapter, operating over billions of years, have produced a world in which life is organized into millions of species, each adapted to live in a particular environment and to use environmental resources in a particular way. In the next chapter we consider how species evolve at the level of their genes and genomes.

CHAPTER SUMMARY

23.1 What Are Species?

- **Speciation** is the process by which one species splits into two or more daughter species, which thereafter evolve as distinct lineages.
- The **biological species concept** distinguishes species on the basis of **reproductive isolation**.
- The **morphological species concept** distinguishes species on the basis of physical similarities; it often underestimates or overestimates the actual number of reproductively isolated species.
- The **lineage species concept** recognizes evolutionarily independent lineages as species, allowing biologists to consider species over evolutionary time.

23.2 How Do New Species Arise?

- Speciation usually results from the interruption of gene flow within a population.
- The Dobzhansky-Muller model describes how reproductive isolation can develop between two descendant species. [Review Figure 23.2](#)
- **Allopatric speciation**, which results when populations are separated by a physical barrier, is the dominant mode of speciation. This type of speciation may follow from founder events, in which some members of a population cross a barrier and found a new, isolated population. [Review Figure 23.5, ANIMATED TUTORIAL 23.1](#)

- **Sympatric speciation** results when the genomes of two groups diverge in the absence of physical isolation. It can result from disruptive selection for two or more distinct microhabitats.
- Sympatric speciation can occur within two generations via **polyploidy**, an increase in the number of chromosomes sets. Polyploidy may arise from chromosome duplications within a species (**autopolyploidy**) or from hybridization that results in combining the chromosomes of two species (**allopolyploidy**). [Review Figure 23.8](#)

[SEE ANIMATED TUTORIAL 23.2](#)

23.3 What Happens When Newly Formed Species Come Together?

- **Prezygotic barriers** to reproduction operate before fertilization; **postzygotic barriers** to reproduction operate after fertilization. Prezygotic barriers may be favored by natural selection if postzygotic barriers are incomplete. [Review Figures 23.9, 23.11](#)
- **Hybrid zones** may form when previously separated populations come into contact and reproductive isolation is incomplete. [Review Figure 23.14](#)

23.4 Why Do Rates of Speciation Vary?

- Dispersal ability, dietary specialization, type of pollination, and sexual selection all influence speciation rates. [Review Figure 23.15](#)

[SEE WEB ACTIVITY 23.1 for a concept review of this chapter.](#)

SELF-QUIZ

1. The biological species concept defines a species as a group of
 - a. actually interbreeding natural populations that are reproductively isolated from other such groups.
 - b. potentially interbreeding natural populations that are reproductively isolated from other such groups.
 - c. actually or potentially interbreeding natural populations that are reproductively isolated from other such groups.
 - d. actually or potentially interbreeding natural populations that are reproductively connected to other such groups.
 - e. actually interbreeding natural populations that are reproductively connected to other such groups.
2. Which of the following is *not* a condition expected to favor allopatric speciation?
 - a. Continents drift apart and separate previously connected lineages.
 - b. A mountain range separates formerly connected populations.
 - c. Different environments on two sides of a barrier cause populations to diverge.
 - d. The range of a species is separated by loss of intermediate habitat.
 - e. Tetraploid individuals arise in one part of the range of a species.
3. Finches speciated in the Galápagos Islands because
 - a. the Galápagos Islands are not far from the mainland.
 - b. the Galápagos Islands are thought to promote sympatric speciation in birds.
 - c. hybridization across different island populations of finches led to high levels of polyploidy.
4. Which of the following is *not* a potential prezygotic reproductive barrier?
 - a. Temporal segregation of breeding seasons
 - b. Differences in chemicals that attract mates
 - c. Hybrid infertility
 - d. Spatial segregation of mating sites
 - e. Sperm that cannot penetrate an egg
5. A common means of sympatric speciation is
 - a. polyploidy.
 - b. hybrid infertility.
 - c. temporal segregation of breeding seasons.
 - d. spatial segregation of mating sites.
 - e. imposition of a geographic barrier.
6. Narrow hybrid zones may persist for long times because
 - a. hybrids are always at a disadvantage.
 - b. hybrids have an advantage only in narrow zones.
 - c. hybrid individuals never move far from their birthplaces.
 - d. individuals that move into the zone have not previously encountered individuals of the other species, so reinforcement of reproductive barriers has not occurred.
 - e. Narrow hybrid zones are artifacts because biologists generally restrict their studies to contact zones between species.

7. Which statement about speciation is *not* true?
 - a. It always takes thousands of years.
 - b. Reproductive isolation may develop slowly between diverging lineages.
 - c. Among animals, it usually requires a physical barrier.
 - d. Among plants, it often happens as a result of polyploidy.
 - e. It has produced the millions of species living today.
8. Which of the following is often associated with higher rates of speciation?
 - a. Sexually dimorphic compared with sexually monomorphic birds
 - b. Insects with specialized diets compared with insects with generalized diets
 - c. Species with low dispersal ability compared with species with high dispersal ability
 - d. Plants with animal pollination compared with plants with wind pollination
 - e. All of the above
9. Evolutionary radiations
 - a. happen often on continents but rarely on island archipelagoes.
 - b. characterize birds and plants but not other groups of organisms.
 - c. have happened on continents as well as on islands.
 - d. require major reorganizations of the genome.
 - e. never happen in species-poor environments.
10. Speciation is an important component of evolution because it
 - a. generates the variation on which natural selection acts.
 - b. generates the variation on which genetic drift and mutations act.
 - c. enabled Charles Darwin to perceive the mechanisms of evolution.
 - d. generates the high extinction rates that drive evolutionary change.
 - e. has resulted in a world with millions of species, each adapted for a particular way of life.

FOR DISCUSSION

1. The North American snow goose has two distinct color forms, blue and white. Matings between the two color forms are common. However, blue individuals pair with blue individuals and white individuals pair with white individuals much more frequently than would be expected by chance. Suppose that blue and white snow geese are equally frequent in a population, and that 75 percent of all mated pairs consist of two individuals of the same color. What would you conclude about speciation processes in these geese? What if 100 percent of pairs were the same color?
2. Suppose pairs of snow geese of mixed colors were found only in a narrow zone within the broad Arctic breeding range of the geese, with blue geese found on one side and white geese found on the other side of this narrow zone. Would your answer to Question 1 remain the same?
3. Although many butterfly species are divided into local populations among which there is little gene flow, these species often show relatively little morphological variation among populations. Describe the studies you would conduct to determine what maintains this morphological similarity.
4. Evolutionary radiations are common and easily studied on oceanic islands. In what types of *mainland* situations would you expect to find major evolutionary radiations? Why?
5. Fruit flies of the genus *Drosophila* are distributed worldwide, but 30 to 40 percent of all the species in the genus are found on the Hawaiian Islands (which comprise far less than 1% of Earth's total land area). What might account for this distribution pattern?
6. What factors can cause extinction rates to exceed speciation rates in a clade? Name some clades in which human activities are increasing extinction rates without increasing speciation rates.
7. If it is true that natural selection does not directly favor lower viability of hybrids, why is it that hybrid individuals so often have lowered viability?

ADDITIONAL INVESTIGATION

In the two *Aquilegia* species shown in Figure 23.12, the orientation of the flowers and the length of flower spurs are associated with the respective pollinator species (hummingbirds and hawkmoths). Columbine flowers vary in other ways as well; for

example, they differ in color, and probably in odor. What experiments could you design to determine the traits that various pollinators use to distinguish among the flowers of different columbine species?

WORKING WITH DATA (GO TO yourBioPortal.com)

Examining Evidence for Reinforcement of Prezygotic Barriers

In this exercise based on Figure 23.13, you will examine some of the data collected by Don Levin to study reinforcement of

prezygotic reproductive barriers in *Phlox*. You will also critique the study design of the experiment, and consider alternative explanations for the results.

Shocking evolution

Some fishes, including the famous electric eel of Central and South America, can produce high-voltage discharges of electricity (up to 650 volts) that they use to stun their prey. A variety of other fish species are known to produce somewhat weaker electric discharges. Most of these latter species live in murky water where visual cues are limited; they use electric signals to locate (but not to stun) their prey. Electric signals also allow them to communicate with other individuals of their own species.

Electric organs have evolved independently in several fish lineages. How did these organs evolve? Let's consider first the physical basis of the electrical signal. *Voltage-gated sodium channels* are large proteins that underlie the generation and propagation of rapid electrical signals in nerve, muscle, and heart tissues (see Chapter 45). Electric signals are transmitted along nerves to muscles as the sodium channels embedded in cell membranes are stimulated to open. These channels control the concentration of positively charged sodium ions (Na^+) on the inside relative to the outside of cells, resulting in an electric charge

that is transmitted across the surface of the muscle, leading to muscular contraction.

Most vertebrates have a number of different copies of the genes encoding the several proteins that make up the sodium channel. These copies arose through a series of *gene duplications* in the distant past of vertebrate genome evolution. Such duplications allowed for the specialization of protein function, making it possible for different sodium channels to exist in different types of tissue. In the case of electric fishes, one of the sodium channel genes ordinarily expressed in muscle diverged and a new functional protein evolved. Changes in a relatively small number of nucleotide positions in the gene resulted in modified sodium channels, allowing the development of a new organ with a unique function—the generation of externally transmitted electric energy.

The “living battery” electric organ differs from skeletal muscle in important ways. The organ is composed of many *electrocytes*, each of which is a derived muscle cell capable of producing a small electric charge. Electrocytes

are stacked in series, much like the plates in a car battery. Rather than producing muscle contraction and movement, however, the organ generates an electric discharge. This signal is species-specific, which allows intraspecific communication and also serves as an isolating mechanism between species (see Chapter 23).

The repeated evolution of electric organs from muscle tissue is facilitated by relatively simple molecular changes



An Electric Fish The elephant-nose fish (*Gnathonemus petersi*), a river-dwelling species from West Africa, is one of many fishes in which weakly discharging electric organs have evolved via modifications in sodium channel proteins.



A High-Voltage Electric Fish This torpedo ray can put out as much as 220 volts of electricity. So far this particular species remains unidentified; it has been found only in a single bay among the islands of Komodo National Park, Indonesia.

in certain genes, changes that result in major functional changes to sodium channels. Gene duplication facilitates the process, since “extra” genes allow for such specialization in protein function. Finally, interspecific differences in sodium channel function result from additional changes in nucleotide sequences of the respective genes. These small differences allow different species of fishes to use different communication signals, which improves intraspecific communication while reducing interspecific interference.

The evolution of sodium channels is just one example of how an understanding of the evolution of genes and genomes helps biologists understand the diversity of life on Earth. Molecular investigations also allow biologists to observe the process of evolution directly in the laboratory, and to use evolutionary principles to produce new molecules with useful functions.

IN THIS CHAPTER we will see how molecular biologists infer both the patterns and the causes of molecular evolution from studies of nucleic acids and proteins. We will explore how the functions of molecules change, how genomes change in size, and where new genes come from. Finally, we will explore some practical applications of molecular evolution for producing new molecules with novel functions.

CHAPTER OUTLINE

- 24.1** How Are Genomes Used to Study Evolution?
- 24.2** What Do Genomes Reveal About Evolutionary Processes?
- 24.3** How Do Genomes Gain and Maintain Functions?
- 24.4** What Are Some Applications of Molecular Evolution?

24.1 How Are Genomes Used to Study Evolution?

An organism’s **genome** is the full set of genes it contains, as well as any noncoding regions of the DNA (or in the case of some viruses, RNA). Most of the genes of eukaryotic organisms are found on chromosomes in the nucleus, but genes are also present in chloroplasts and mitochondria. In organisms that reproduce sexually, both males and females transmit nuclear genes, but mitochondrial and chloroplast genes usually are transmitted only via the cytoplasm of one of the two gametes (usually from the female parent).

Genomes must be replicated to be transmitted from parents to offspring. DNA replication does not occur without error, however. Mistakes in DNA replication—mutations—provide much of the raw material for evolutionary change. Mutations are essential for the long-term survival of life, because they are the initial source of the genetic variation that permits organisms to evolve in response to changes in their environment.

A particular copy of a gene will not be passed on to successive generations unless an individual with that copy survives and reproduces. Therefore, the capacity to cooperate with different combinations of other genes is likely to increase the probability that a particular allele will become fixed in a population. Moreover, the degree and timing of a gene’s expression are affected by its location in the genome. For these reasons, the genes of an individual organism can be viewed as interacting members of a group, among which there are divisions of labor but also strong interdependencies.

A genome, then, is not simply a random collection of genes in random order along chromosomes. Rather, it is a complex set of integrated genes, regulatory sequences, and structural elements, as well as vast stretches of noncoding DNA that may have little direct function. The positions of genes, as well as their sequences, are subject to evolutionary change, as are the extent and location of noncoding DNA. All of these changes can affect the phenotype of an organism. Biologists have now sequenced the complete genomes of a large number of organisms, including humans. This information is helping us to understand how and why organisms differ, how they function, and how they have evolved.

Evolution of genomes results in biological diversity

The field of **molecular evolution** investigates the mechanisms and consequences of the evolution of macromolecules. Molecular evolutionists study relationships between the structures of genes and proteins and the functions of organisms. They also

examine molecular variation to reconstruct evolutionary history and to study the mechanisms and consequences of evolution. The molecules of special interest to molecular evolutionists are nucleic acids (DNA and RNA) and proteins. Students of this field ask questions such as: What does molecular variation tell us about a gene's function? Why do the genomes of different organisms vary in size? What evolutionary forces shape patterns of variation among genomes? And a crucial question from an evolutionary perspective, How do genomes acquire new functions? Investigations into the evolution of particular nucleic acids and proteins are instrumental in reconstructing the evolutionary histories of genes and in determining which organisms carry them. Ultimately, molecular evolutionists hope to explain the molecular basis of biological diversity.

The evolution of nucleic acids and proteins depends on genetic variation introduced by mutations. One of several ways in which genes evolve is by means of *nucleotide substitutions*. In genes that encode proteins, nucleotide substitutions sometimes result in amino acid replacements that can change the charge, the structure (secondary or tertiary), and other chemical and physical properties of the encoded protein. Phenotypic changes in the protein often affect the way that protein functions in the organism.

Evolutionary changes in genes and proteins can be identified by comparing the nucleotide or amino acid sequences of different organisms. The longer two sequences have been evolving separately, the more differences they accumulate (bearing in mind that different genes in the same species evolve at different rates). Determining when changes in nucleotide or amino acid sequences occurred is a first step toward inferring their causes. Knowledge of the pattern and rate of evolutionary change in a given macromolecule is useful in reconstructing the evolutionary history of groups of organisms.

To compare genes or proteins across different organisms, biologists need a way to identify homologous parts of molecules. (Recall from Section 22.1 that *homologous* features are shared by two or more species and have been inherited from a common ancestor.) Homologous parts of a protein can be traced to homologous amino acid sequences. And, since nucleotide sequences encode amino acid sequences, the concept of homology extends down to the level of individual nucleotide positions. Therefore, one of the first steps in studying the evolution of genes or proteins is to align homologous positions in the nucleotide or amino acid sequence of interest.

Genes and proteins are compared through sequence alignment

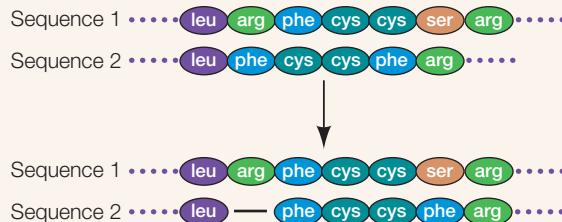
Once the DNA or amino acid sequences of molecules from different organisms have been determined, they can be compared. Homologous positions can be identified only if we first pinpoint the locations of deletions

TOOLS FOR INVESTIGATING LIFE

24.1 Amino Acid Sequence Alignment

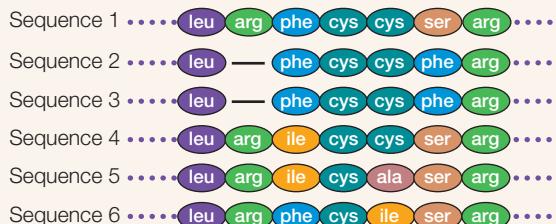
Amino acid sequence alignment is a way of arranging protein sequences to identify regions of homology between the sequences. Gaps are inserted between the amino acid residues to align similar residues in columns. Differences and similarities between each pair of aligned sequences are then summarized in a similarity matrix. Homologous DNA sequences can be aligned in a similar manner.

1 Two amino acid sequences seem quite different...



2 ...but if we insert a gap in Sequence 2, there is nearly complete alignment.

3 With this alignment established, we can compare additional sequences.



Numbers above the diagonal line are the number of differences.

Similarity matrix

		Sequence number					
		1	2	3	4	5	6
Sequence number	1	5	2	2	1	2	1
	2	5	7	0	3	4	3
	3	5	7	3	4	3	
	4	6	4	4		1	2
	5	5	3	3	6		2
	6	6	4	4	5	5	

Numbers below the diagonal line are similarities.

yourBioPortal.com

GO TO Web Activity 24.1 • Amino Acid Sequence Alignment

and insertions that have occurred in the molecules of interest in the time since the organisms diverged from a common ancestor. A simple hypothetical example illustrates this **sequence alignment** technique. In **Figure 24.1** we compare two amino acid sequences (1 and 2) from homologous proteins in different organisms. The two sequences at first appear to differ in both the number and identity of their amino acids, but if we insert a gap after the first amino acid in Sequence 2 (after leucine), similarities in the two sequences become more obvious. This gap represents the occurrence of one of two evolutionary events: an insertion of an amino acid in the longer protein, or a deletion of an amino acid in the shorter protein.

Having adjusted for this insertion or deletion event, we can see that the two sequences differ by only one amino acid at position 6 (serine or phenylalanine). Adding a single gap—that is, identifying a deletion or an insertion—*aligns* these sequences. Longer sequences and those that have diverged more extensively require more elaborate adjustments. Explicit models (incorporated into computer algorithms) have been developed to account for the relative probabilities of deletions, insertions, and particular amino acid replacements.

Having aligned the sequences, we can compare them by counting the number of nucleotides or amino acids that differ between them. Summing the number of similar and different amino acids in each pair of sequences allows us to construct a **similarity matrix**, which gives us a measure of the minimum number of changes that have occurred during the divergence between each pair of organisms (see Figure 24.1).

yourBioPortal.com

GO TO Web Activity 24.2 • Similarity Matrix Construction

Models of sequence evolution are used to calculate evolutionary divergence

The sequence comparison procedure illustrated in Figure 24.1 gives a simple count of the number of differences and similarities between the proteins of two species. In the context of two aligned DNA sequences, we can count the number of differences at homologous nucleotide positions, and this count indicates the minimum number of nucleotide changes that must have occurred since the two sequences diverged from a common ancestral sequence.

Although it is useful in determining a *minimum* number of changes between two sequences, the count provided by sequence alignment almost certainly underestimates the actual number of changes that have occurred since the sequences diverged. Any given change counted in a similarity matrix of DNA sequences may result from multiple substitution events

24.2 Multiple Substitutions Are Not Reflected in Pairwise Sequence Comparisons

Two observed sequences are descended from a common ancestral sequence (center) through a series of substitutions. Although the two observed sequences differ in only three nucleotide differences (colored letters), these three differences resulted from a total of nine substitutions (arrows).

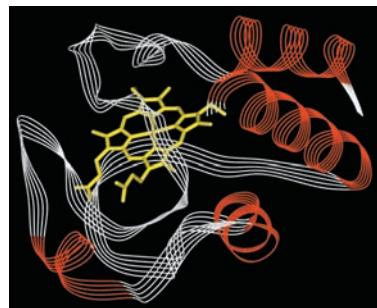
that occurred at a given nucleotide position over time. As illustrated in **Figure 24.2**, any of the following events may have occurred at a given nucleotide position that would not be revealed by a simple count of similarities and differences between two DNA sequences:

- **Multiple substitutions.** More than one change occurs at a given position between the ancestral sequence and at least one of the observed sequences.
- **Coincident substitutions.** At a given position, different substitutions occur between the ancestral sequence and each observed sequence.
- **Parallel substitutions.** The same substitution occurs independently between the ancestral sequence and each observed sequence.
- **Back substitutions** (also called *reversions*). In a variation on multiple substitutions, after a change at a given position, a subsequent substitution changes the position back to the ancestral state.

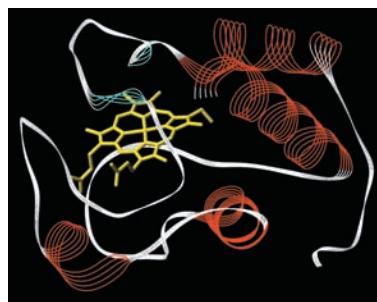
To correct for undercounting of substitutions, molecular evolutionists have developed mathematical models that describe how DNA (and protein) sequences evolve. These models take into account the relative rates of change from one nucleotide to another; for example, *transitions* (changes between the two purines, A \leftrightarrow G, or between the pyrimidines, C \leftrightarrow T) are more frequent than *transversions* (a purine is replaced by a pyrimidine, or vice versa). Models also include parameters such as the different rates of substitution across different parts of a gene and the proportions of each nucleotide present in a given sequence. Once such parameters have been estimated, the model

Observed sequence 1	Ancestral sequence	Observed sequence 2	
A	A	A	
C	C	T	Single substitution
G	G	G	
G	G	G	
T	T	C	Multiple substitutions
A	A	A	
T	T	T	
T	T	T	
G	A	T	Coincident substitutions
C	C	C	
C	G	C	Parallel substitutions
G	G	G	
G	G	G	
C	C	C	
T	T	T	
A	A	G	* Back substitution (reversion)
T	T	A	
A	A	A	

Tuna



Rice

**Acidic side chains**

D Aspartic acid
E Glutamic acid

Basic side chains

H Histidine
K Lysine
R Arginine

Hydrophobic side chains

F Phenylalanine
I Isoleucine
L Leucine
M Methionine

V Valine
Y Tyrosine
W Tryptophan
A Alanine

Cysteine
Proline
Glutamine
Asparagine
Serine
Threonine
Glycine

The number 1 indicates an invariant position in the cytochrome c molecule (i.e., all the organisms have the same amino acid in this position). Such a position is probably under strong purifying selection.

Amino acids at positions marked by red arrowheads have side chains that interact with the heme group.

Position in sequence	1	5	10	15	20	25	30
Number of amino acids at the position	1 3 5 5 5	1 3 3 4	1 4 3 2	1 3 3 1	1 2 4 3	2 3 4 2	2 1 4 1 1 2 1 5 1
Human, chimpanzee	G D V E K G K K	I F I M K C S Q C H	T V E K G G K H K T G P N L	H G			
Rhesus monkey	G D V E K G K K	I F I M K C S Q C H	T V E K G G K H K T G P N L	H G			
Horse	G D V E K G K K	I F V Q K C A Q C H	T V E K G G K H K T G P N L	H G			
Donkey	G D V E K G K K	I F V Q K C A Q C H	T V E K G G K H K T G P N L	H G			
Cow, pig, sheep	G D V E K G K K	I F V Q K C A Q C H	T V E K G G K H K T G P N L	H G			
Dog	G D V E K G K K	I F V Q K C A Q C H	T V E K G G K H K T G P N L	H G			
Rabbit	G D V E K G K K	I F V Q K C A Q C H	T V E K G G K H K T G P N L	H G			
Gray whale	G D V E K G K K	I F V Q K C A Q C H	T V E K G G K H K T G P N L	H G			
Gray kangaroo	G D V E K G K K	I F V Q K C A Q C H	T V E K G G K H K T G P N L N G				
Chicken, turkey	G D I E K G K K	I F V Q K C S Q C H	T V E K G G K H K T G P N L	H G			
Pigeon	G D I E K G K K	I F V Q K C S Q C H	T V E K G G K H K T G P N L	H G			
Pekin duck	G D V E K G K K	I F V Q K C S Q C H	T V E K G G K H K T G P N L	H G			
Snapping turtle	G D V E K G K K	I F V Q K C A Q C H	T V E K G G K H K T G P N L N G				
Rattlesnake	G D V E K G K K	I F T M K C S Q C H	T V E K G G K H K T G P N L	H G			
Bullfrog	G D V E K G K K	I F V Q K C A Q C H	T C E K G G K H K V G P N L Y G				
Tuna	G D V A K G K K	T F V Q K C A Q C H	T V E N G G K H K V G P N L W G				
Dogfish	G D V E K G K K	V F V Q K C A Q C H	T V E N G G K H K T G P N L S G				
Samia cynthia (moth)	G N A E N G K K	I F V Q R C A Q C H	T V E A G G K H K V G P N L	H G			
Tobacco hornworm moth	G N A D N G K K	I F V Q R C A Q C H	T V E A G G K H K V G P N L	H G			
Screwworm fly	G D V E K G K K	I F V Q R C A Q C H	T V E A G G K H K V G P N L	H G			
Drosophila (fruit fly)	G D V E K G K K	L F V Q R C A Q C H	T V E A G G K H K V G P N L	H G			
Baker's yeast	G S A K K G A T L	F K T R C A E C H	T V E A G G K H K V G P N L	H G			
Candida krusei (yeast)	G S A K K G A T L	F K T R C A E C H	T V E A G G K H K V G P N L	H G			
Neurospora crassa (mold)	G D S K K G A N L	F K T R C A E C H	— E — N L T Q K I G P A L H G				
Wheat	G N P D A G A K I	F K T K C A Q C H	T V D A G A — H K Q G P N L H G				
Sunflower	G D P T T G A K I	F K T K C A Q C H	T V E K G A — H K Q G P N L N G				
Mung bean	G D S K S G E K I	F K T K C A Q C H	T V D K G A — H K Q G P N L N G				
Rice	G N P K A G E K I	F K T K C A Q C H	T V D K G A — H K Q G P N L N G				
Sesame	G D V K S G E K I	F K T K C A Q C H	T V D K G A — H K Q G P N L N G				

Gaps indicate insertion and/or deletion events.

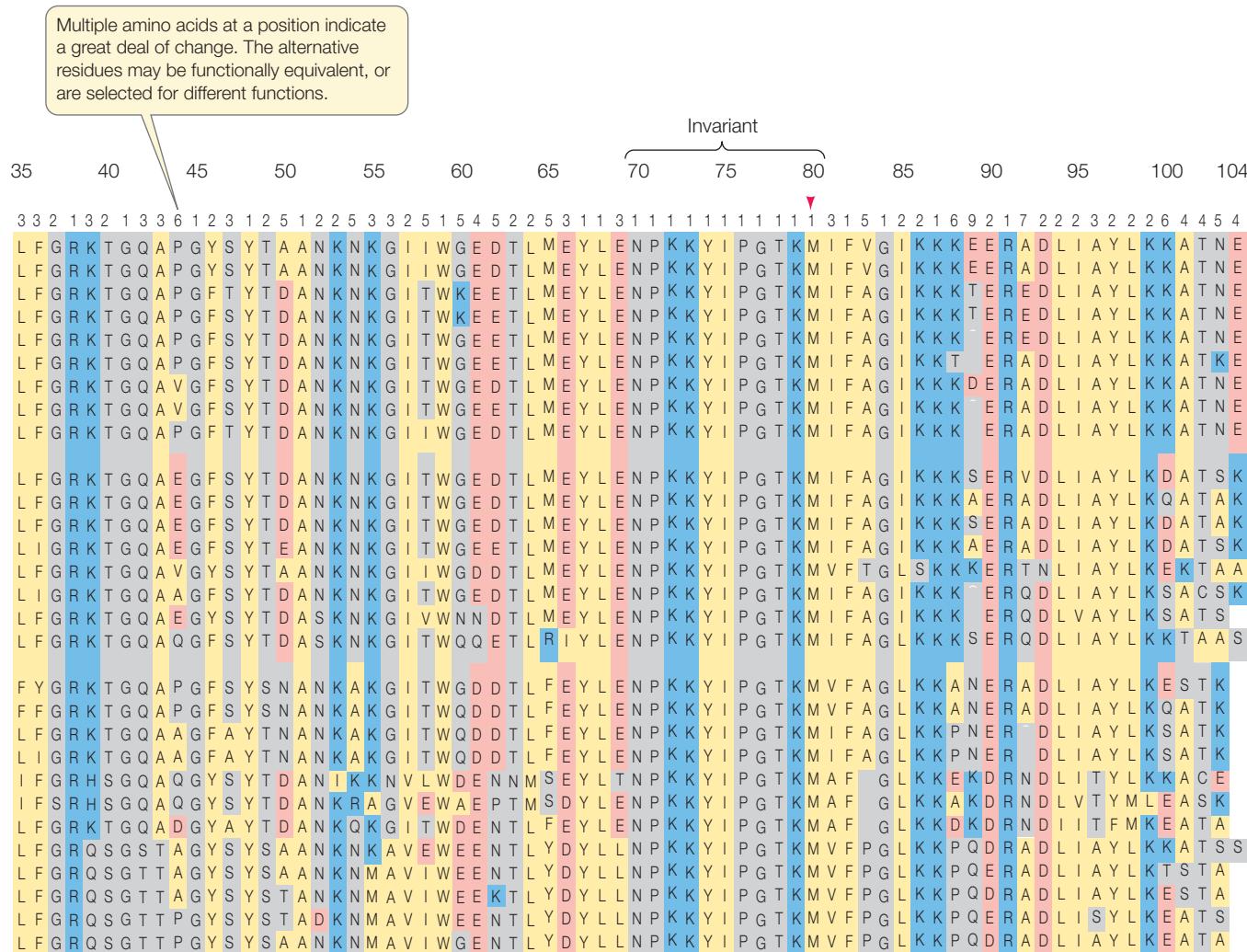
24.3 Amino Acid Sequences of Cytochrome c The amino acid sequences shown in the table were obtained from analyses of the enzyme cytochrome c from 33 species of plants, fungi, and animals. Note the lack of variation across the sequences at positions 70–80, suggesting that this region is under strong stabilizing selection and that changing its amino acid sequence would impair the protein's function. The computer graphics at the upper left are created from these sequences and show the three-dimensional structures of tuna and rice cytochrome c. Alpha helices are in red, and the molecule's heme group is shown in yellow.

is used to correct for multiple substitutions, coincident substitutions, parallel substitutions, and back substitutions. The revised estimate accounts for the *total* number of substitutions likely to have occurred between two sequences, which is almost always greater than the observed number of differences.

As sequence information becomes available for more and more genes in an ever-expanding database, sequence alignments can be extended across multiple homologous sequences, and the minimum number of insertions, deletions, and substitutions can be summed across homologous genes of an entire group of organisms. Similar databases have also been constructed for homologous proteins. **Figure 24.3** shows aligned data for cytochrome c protein sequences in 33 species of animals, plants, and fungi. Such information is used extensively in determining evolutionary relationships among species.

Experimental studies examine molecular evolution directly

Although molecular evolutionists are often interested in naturally evolved genes and proteins, molecular and phenotypic evolution can also be observed directly in the laboratory. Increasingly, evolutionary biologists are studying evolution experimentally. Because substitution rates are related to generation time rather than to absolute time, most of these experiments use unicellular organisms or viruses with short generations. Viruses, bacteria, and unicellular eukaryotes (such as the yeasts) can be cultured in large populations in the laboratory, and many of these organisms can evolve rapidly. In the case of some RNA viruses, the natural substitution rate may be as high as 1 substitution per 1,000 nucleotides per generation. Therefore, in a virus of a few thousand nucleotides, one



or more substitutions are expected (on average) every generation, and these changes can easily be determined by sequencing the entire genome (because of its small size). Generation time may be only tens of minutes (rather than years or decades, as in humans), so biologists can directly observe substantial molecular evolution in a controlled population over the course of days, weeks, or months.

An example of an experimental evolutionary study is shown in **Figure 24.4**. Paul Rainey and Michael Travisano wanted to examine a potential cause of adaptive radiations, which are a major source of biological diversity (see Section 23.4). For instance, near the beginning of the Cenozoic era, mammals rapidly diversified into species as diverse as elephants, moles, whales, and bats. While Rainey and Travisano clearly couldn't experimentally manipulate mammals over many millions of years, they could test the idea that heterogeneous environments with unoccupied niches lead to adaptive radiation by experimentally manipulating a bacterial lineage.

Rainey and Travisano inoculated several flasks containing culture medium with the same strain of the bacterium *Pseudomonas fluorescens*. They then shook some of the cultures

to maintain a constantly uniform environment. Others they left alone (static cultures), allowing them to develop a spatially distinct structure. In the static cultures, the environment on the surface film of the medium differed from that on the walls of the flasks and from parts of the culture not touching any surfaces.

When the cultures were started, the ancestral phenotype of the bacterium produced a smooth colony, which the investigators called a "smooth morph." In just a few days, however, the static cultures consistently and independently developed two other morphs: a "wrinkly spreader" and a "fuzzy spreader." The researchers determined that the two new morphs had a genetic basis and were adaptively superior in some of the environments found within the static cultures. For example, the "wrinkly spreader" cells adhered firmly to one another as well as to surfaces, and thus were able to form a mat across the surface of the medium, where they could compete successfully for oxygen.

DNA sequencing of the genomes of these morphs showed that the same phenotypes had evolved repeatedly, and that many different substitutions could produce the same phenotypes. The homogeneous shaken cultures, in contrast, showed no evolution in phenotype. The same mutations occurred in the

INVESTIGATING LIFE

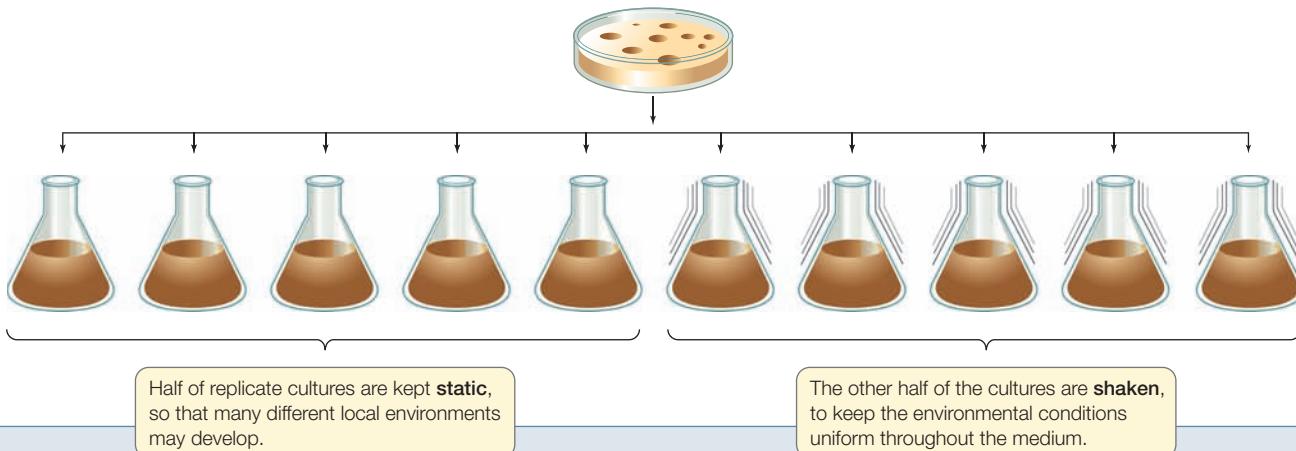
24.4 Evolution in a Heterogeneous Environment

Rainey and Travisano cultured the rapidly evolving bacterium *Pseudomonas fluorescens* in homogeneous and heterogeneous environments to examine the relationship between phenotypic diversity and environmental variability.

HYPOTHESIS Heterogeneous environments are more conducive to the evolution of phenotypic diversity than are homogeneous environments.

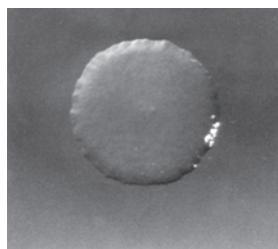
METHOD

One colony of *Pseudomonas fluorescens* (all of a single genotype) is used to inoculate many replicate cultures.

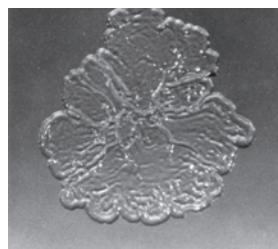


RESULTS

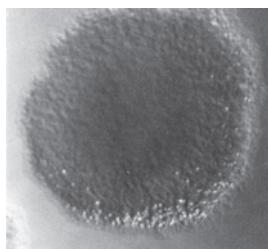
In the shaken flasks, the ancestral morph persisted; the uniform environment did not result in morphological diversification. In the static flasks, two new morphotypes regularly arose, each adapted to a different local environment. Molecular analysis revealed that the mutations that produce these phenotypes arose in both shaken and static cultures, but the mutations did not persist in the uniform (shaken) environment because the phenotypes they produced were selectively disadvantageous under homogeneous conditions.



Smooth morph (ancestral)



“Wrinkly spreader”



“Fuzzy spreader”

CONCLUSION Phenotypic change and diversification are enhanced in a heterogeneous environment.

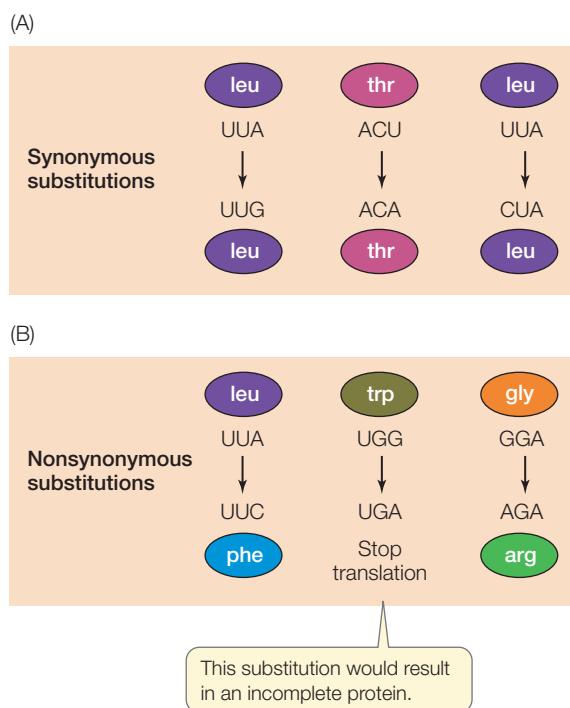
FURTHER INVESTIGATION: Do you think the two evolved phenotypes could compete in the homogeneous environment if they were introduced after having become successfully established in the heterogeneous environment? How would you test your hypothesis?

Go to yourBioPortal.com for original citations, discussions, and relevant links for INVESTIGATING LIFE figures.

shaken cultures but did not persist, because the novel phenotypes they produced were selectively disadvantageous (i.e., less fit) under the “shaken” environmental conditions.

Experimental molecular evolutionary studies are used for a wide variety of purposes and have greatly expanded the abil-

ity of evolutionary biologists to test evolutionary concepts and principles. Biologists now routinely study evolution in the laboratory and, as we will see later in this chapter, use in vitro evolutionary techniques to produce novel molecules that perform new functions with industrial and pharmaceutical uses.



24.5 When One Nucleotide Does or Doesn't Make a Difference

(A) Synonymous substitutions do not change the amino acid specified and do not affect protein function; such substitutions are less likely to be subject to natural selection. (B) Nonsynonymous substitutions do change the amino acid sequence and are likely to have an effect (often deleterious) on protein function; such substitutions are targets for natural selection.

24.2 What Do Genomes Reveal About Evolutionary Processes?

A *mutation*, as we saw in Chapter 15, is any change in the genetic material. A nucleotide substitution is one type of mutation. Many nucleotide substitutions have no effect on phenotype, even if the change occurs in a gene that encodes a protein, because most amino acids are specified by more than one codon (see Figure 14.6). A substitution that does not change the encoded amino acid is known as a **synonymous substitution** or **silent substitution** (Figure 24.5A). Synonymous substitutions do not affect the functioning of a protein (although they may have other effects, such as changes in mRNA stability or translation rates; see Section 14.5), and are therefore less likely to be influenced by natural selection.

A nucleotide substitution that *does* change the amino acid sequence encoded by a gene is known as a **nonsynonymous substitution**, also known as a *missense substitution* (Figure 24.5B). In general, nonsynonymous substitutions are more likely to be deleterious to the organism. But not every amino acid replacement alters a protein's shape and charge (and hence its functional properties). Therefore, some nonsynonymous substitutions may also be selectively neutral, or nearly so. Conversely, an amino acid replacement that confers an advantage to the organism would result in positive selection for the corresponding nonsynonymous substitution.

The rate of nonsynonymous nucleotide substitutions in several mammalian protein-coding genes is about 3×10^{-9} substitutions per position per year. Synonymous substitutions in these genes have occurred about five times more frequently than nonsynonymous substitutions. In other words, *substitution rates are highest at nucleotide positions that do not change the amino acid being expressed* (Figure 24.6). The rate of substitution is even higher in **pseudogenes**, which are duplicate copies of genes that are no longer functional.

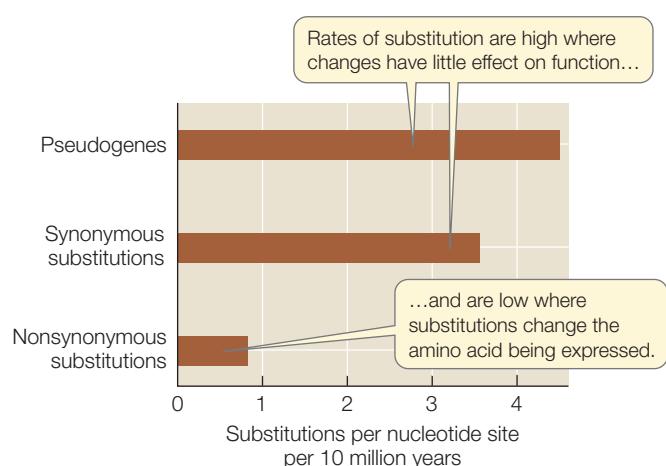
As we saw Chapter 21, most natural populations harbor far more genetic variation than we would expect to find if genetic

24.1 RECAP

The genomes of all organisms evolve over time, as can be detected by direct observation in the laboratory, as well as by aligning genes and proteins between species. Experimental studies of molecular evolution allow biologists to study many processes of evolution directly under controlled conditions.

- How do biologists align nucleotide and amino acid sequences they wish to compare, and how do they calculate the minimum number of changes that have occurred between pairs of aligned sequences? **See pp. 500–501 and Figure 24.1**
- Explain why a simple count of nucleotide differences between two sequences underestimates the actual number of nucleotide substitutions since the sequences diverged. **See Figure 24.2**

We have seen that molecular evolutionists can directly observe the evolution of genomes over time, and can compare the genomes of different organisms and reconstruct the changes that have occurred during their evolution. Let's turn now to the question of how genomes change and examine some of the consequences of those changes.



24.6 Rates of Substitution Differ Rates of nonsynonymous substitution typically are much lower than rates of synonymous substitution and the substitution rate in pseudogenes. This pattern reflects differing levels of functional constraints.

variation were influenced by natural selection alone. This discovery, combined with the knowledge that many mutations do not change molecular function, stimulated the development of the *neutral theory* of molecular evolution.

Much of evolution is neutral

In 1968, Motoo Kimura proposed the *neutral theory of molecular evolution*. Kimura suggested that, at the molecular level, the majority of variants we observe in most populations are selectively neutral; that is, they confer neither an advantage nor a disadvantage on their bearers. Therefore, these neutral variants accumulate through genetic drift rather than through positive selection.

The rate of fixation of neutral mutations by genetic drift is independent of population size. To see why this is so, consider a population of size N and a neutral mutation rate μ (mu) per gamete per generation at a locus. The number of new mutations would be, on average, $\mu \times 2N$, because $2N$ gene copies are available to mutate in a population of diploid organisms. The probability that a given mutation will be fixed by drift alone is its frequency, which equals $1/(2N)$ for a newly arisen mutation. We can multiply these two terms to get the rate of fixation of neutral mutations in a given population of N individuals:

$$2N\mu \frac{1}{2N} = \mu$$

Therefore, the rate of fixation of neutral mutations depends only on the neutral mutation rate μ and is independent of population size. A given mutation is more likely to appear in a large population than in a small one, but any mutation that does appear is more likely to become fixed in a small population. These two influences of population size cancel each other out.

Therefore, the rate of fixation of neutral mutations is equal to the mutation rate. As long as the underlying mutation rate is constant, macromolecules evolving in different populations should diverge from one another in neutral changes at a constant rate. Empirically, the rate of evolution of particular genes and proteins is often relatively constant over time, and therefore can be used as a “molecular clock.” As we described in Section 22.3, molecular clocks can be used to calculate evolutionary divergence times between species.

Although much of the genetic variation we observe in populations is the result of neutral evolution, the neutral theory does not imply that most mutations have no effect on the organism. Many mutations are never observed in populations because they are lethal or strongly detrimental to the organism and are thus quickly removed from the population through natural selection. Similarly, mutations that confer a selective advantage tend to be quickly fixed in populations, so they do not result in variation at the population level either. Nonetheless, if we compare homologous proteins from different populations or species, some amino acid positions will remain constant under purifying selection, others will vary through neutral genetic drift, and still others will differ between species as a result of positive selection for change. How can these evolutionary processes be distinguished?

Positive and purifying selection can be detected in the genome

As we have just seen, substitutions in a protein-coding gene can be either synonymous or nonsynonymous, depending on whether they change the resulting amino acid sequence of the protein. The relative rates of synonymous and nonsynonymous substitutions are expected to differ in regions of genes that are evolving neutrally, under positive selection for change, or staying unchanged under purifying selection.

- If a given amino acid in a protein can be one of many alternatives (without changing the protein’s function), then an amino acid replacement is *neutral* with respect to the fitness of an organism. In this case, the rates of synonymous and nonsynonymous substitutions in the corresponding DNA sequences are expected to be very similar, so the ratio of the two rates would be close to 1.
- If a given amino acid position is under *positive selection for change*, the rate of nonsynonymous substitutions is expected to exceed the rate of synonymous substitutions in the corresponding DNA sequences.
- If a given amino acid position is under *purifying selection*, then the rate of synonymous substitutions in the corresponding DNA sequences is expected to be much higher than the rate of nonsynonymous substitutions.

By comparing the gene sequences that encode proteins from many species, scientists can determine the history and timing of synonymous and nonsynonymous substitutions. This information can be mapped on a phylogenetic tree, as we saw in Chapter 22.

Regions of genes that are evolving under neutral, purifying, or positive selection can be identified by comparing the nature and rates of substitutions across the phylogenetic tree. The evolution of lysozyme illustrates how and why particular positions of a gene sequence might be under different modes of selection.

The enzyme lysozyme (see Figure 3.8) is found in almost all animals. It is produced in the tears, saliva, and milk of mammals and in the albumen (whites) of bird eggs. Lysozyme digests the cell walls of bacteria, rupturing and killing them. As a result, it plays an important role as a first line of defense against invading bacteria. Most animals defend themselves against bacteria by digesting them, which is probably why most animals have lysozyme. Some animals also use lysozyme in the digestion of food.

Among mammals, a mode of digestion called *foregut fermentation* has evolved twice. In mammals with this mode of digestion, the foregut—the posterior esophagus and/or the stomach—has been converted into a chamber in which bacteria break down ingested plant matter by fermentation. Foregut fermenters can obtain nutrients from the otherwise indigestible cellulose that makes up a large proportion of plant tissue. Foregut fermentation evolved independently in ruminants (a group of hooved mammals that includes cattle) and in certain leaf-eating monkeys, such as langurs. We know these evolutionary events were independent, because both langurs and ruminants have close relatives that are not foregut fermenters.

In both foregut-fermenting lineages, the enzyme lysozyme has been modified to play a new, nondefensive role. This lysozyme ruptures some of the bacteria that live in the foregut, releasing nutrients metabolized by the bacteria, which the mammal then absorbs. How many changes in the lysozyme molecule were needed to allow it to perform this function amid the digestive enzymes and acidic conditions of the mammalian foregut? To answer this question, molecular evolutionists compared the lysozyme-coding sequences in foregut fermenters with those in several of their nonfermenting relatives. They determined which amino acids differed and which were shared among the species (Figure 24.7A), as well as the rates of synonymous and nonsynonymous substitutions in lysozyme genes across the evolutionary history of the sampled species.

For many of the amino acid positions of lysozyme, the rate of synonymous substitutions (in the corresponding gene) is much higher than the rate of nonsynonymous substitutions. This observation indicates that many of the amino acids that make up lysozyme are evolving under purifying selection. In other words, there is selection against change in the protein at these positions, and the observed amino acids must therefore be critical for lysozyme function. At other positions, several different amino acids function equally well, and the corresponding regions of the genes have similar rates of synonymous and nonsynonymous substitutions. The most striking finding is that amino acid replacements in lysozyme happened at a much higher rate in the lineage leading to langurs than in any other primates. The high rate of nonsynonymous substitutions in the langur lysozyme gene shows that lysozyme went through a period of rapid change in adapting to the stomachs of langurs. Moreover, the lysozymes of langurs and cattle share five amino acid replacements, all of which lie on the surface of the lysozyme molecule, well away from the enzyme's active site. Several of these shared replacements involve changes from arginine to lysine, which makes the proteins more resistant to at-

tack by the stomach enzyme pepsin. By understanding the functional significance of amino acid replacements, molecular evolutionists can explain the observed changes in amino acid sequences in terms of changes in the functioning of the protein.

A large body of fossil, morphological, and molecular evidence shows that langurs and cattle do not share a recent common ancestor. However, langur and ruminant lysozymes share several amino acids that neither mammal shares with the lysozymes of its own closer relatives. The lysozymes of these two mammals have undergone *evolutionary convergence* at some amino acid positions despite their very different ancestry. The amino acids they share give these lysozymes the ability to lyse the bacteria that ferment plant material in the foregut.

The hoatzin, an unusual leaf-eating South American bird and the only known avian foregut fermenter, offers another remarkable story of the convergent evolution of lysozyme (Figure 24.7B). Many birds have an enlarged esophageal chamber called a *crop*. The crop of the hoatzin contains lysozyme and bacteria and acts as a fermenting chamber. Many of the amino acid replacements that occurred in the adaptation of hoatzin crop lysozyme are identical to those that evolved in ruminants and langurs. Thus, even though the hoatzin and foregut-fermenting mammals have not shared a common ancestor in hundreds of millions of years, they have all evolved similar adaptations in their lysozymes that enable them to recover nutrients from their fermenting bacteria.

Genome size and organization also evolve

We know that genome size varies tremendously among organisms. Across broad taxonomic categories, there is some correlation between genome size and organismal complexity. The

24.7 Convergent Molecular Evolution of Lysozyme (A) The number of amino acid differences in the lysozymes of several pairs of mammals are shown above the diagonal line; the number of convergent similarities between these same pairs are shown below the diagonal. The two foregut-fermenting species share convergent amino acid replacements related to this digestive adaptation. (B) The hoatzin—the only known

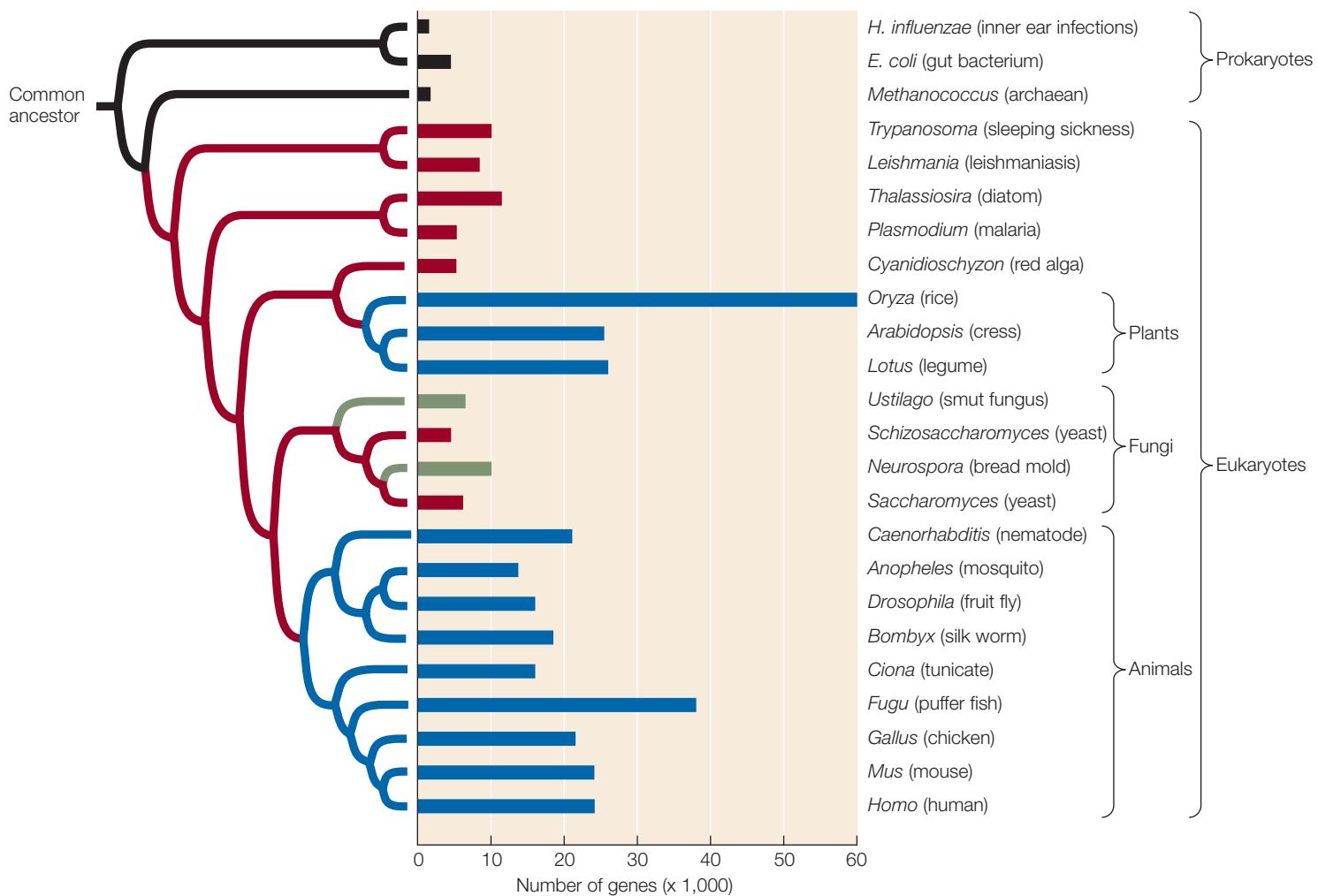
foregut-fermenting bird species—has been evolving independently from mammals for hundreds of millions of years but has independently evolved similar modifications to lysozyme.

(B) *Opisthocomus hoazin*



The lysozymes of langurs and cattle are convergent for 5 amino acid residues, indicative of the independent evolution of foregut fermentation in these two species.

	Langur	Baboon	Human	Rat	Cattle	Horse
Langur		14	18	38	32	65
Baboon	0		14	33	39	65
Human	0	1		37	41	64
Rat	0	0	0		55	64
Cattle	5	0	0	0		71
Horse	0	0	0	0	1	

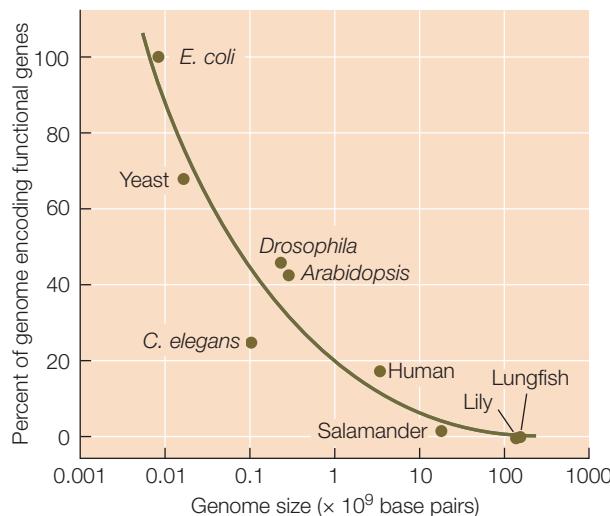


24.8 Genome Size Varies Widely This figure shows the number of genes from a sample of organisms with fully sequenced genomes, arranged by their evolutionary relationships. Bacteria and archaea typically have fewer genes than most eukaryotes. Among eukaryotes, multicellular organisms with tissue organization (plants and animals; blue branches) have more genes than single-celled organisms (red branches) or multicellular organisms that lack pronounced tissue organization (green branches).

genome of the tiny bacterium *Mycoplasma genitalium* has only 470 genes. *Rickettsia prowazekii*, the bacterium that causes typhus, has 634 genes. *Homo sapiens*, by contrast, has about 23,000 protein-coding genes. Figure 24.8 shows the number of genes from a sample of organisms with fully sequenced genomes, arranged by their evolutionary relationships. As this figure reveals, however, a larger genome does not always indicate greater complexity (compare rice to the other plants, for example). It is not surprising that more complex genetic instructions are needed for building and maintaining a large, multicellular organism than a small, single-celled bacterium. What is surprising is that some organisms, such as lungfishes, some salamanders, and lilies, have about 40 times as much DNA as humans do. Structurally, a lungfish or a lily is not 40 times more complex than a human. So why does genome size vary so much?

Differences in genome size are not so great if we take into account only the portion of DNA that actually encodes RNAs or proteins. The organisms with the largest total amounts of nuclear DNA (some ferns and flowering plants) have 80,000 times as much DNA as do the bacteria with the smallest genomes, but no species has more than about 100 times as many protein-coding genes as a bacterium. Therefore, much of the variation in genome size lies not in the number of functional genes but in the amount of noncoding DNA (Figure 24.9).

Why do the cells of most eukaryotic organisms have so much noncoding DNA? Does this noncoding DNA have a function, or is it “junk”? Although some of this DNA does not appear to have a direct function, it can alter the expression of the surrounding genes. The degree or timing of gene expression can be changed dramatically depending on the gene’s position relative to noncoding sequences. Other regions of non-coding DNA consist of pseudogenes that are simply carried in the genome because the cost of doing so is very small. These pseudogenes may become the raw material for the evolution of new genes with novel functions. Some noncoding sequences function in maintaining chromosomal structure. Still others consist of parasitic transposable elements that spread through populations because they reproduce faster than the host genome.



24.9 A Large Proportion of DNA Is Noncoding Most of the DNA of bacteria and yeasts encodes RNAs or proteins, but a large percentage of the DNA of multicellular species is noncoding.

Investigators can use retrotransposons to estimate the rates at which species lose DNA. Retrotransposons are transposable elements (see Figure 17.5) that copy themselves through an RNA intermediate. The most common type of retrotransposon carries duplicated sequences at each end, called long terminal repeats, or LTRs. Occasionally, LTRs recombine in the host genome, so that the DNA between them is excised. When this happens, one recombined LTR is left behind. The number of such “orphaned” LTRs in a genome is a measure of how many retrotransposons have been lost. By comparing the number of LTRs in the genomes of Hawaiian crickets (*Laupala*) and fruit flies (*Drosophila*), investigators found that *Laupala* loses DNA more than 40 times more slowly than does *Drosophila*. Therefore, it is not surprising that the genome of *Laupala* is much larger than that of *Drosophila*.

Why do species differ so greatly in the rate at which they gain or lose apparently functionless DNA? One hypothesis is that genome size is related to the rate at which the organism develops, which may be under selection pressure. Large genomes can slow down the rate of development and thus alter the relative timing of expression of particular genes. As discussed in Section 20.2, changes in the timing of gene expression (*heterochrony*) can produce major changes in phenotype. Thus, although some noncoding DNA sequences may have no direct function, they may still affect the development of the organism.

Another hypothesis is that the proportion of noncoding DNA is related primarily to population size. Noncoding sequences that are only slightly deleterious to the organism are likely to be purged by selection most efficiently in species with large population sizes. In species with small populations, the effects of genetic drift can overwhelm selection against noncoding sequences that have small deleterious consequences. Therefore, selection against the accumulation of noncoding sequences is most effective in species with large populations, so such species (such as bacteria or yeasts) have relatively little noncoding DNA compared with species with small populations (see Figure 24.9).

24.2 RECAP

By examining the relative rates of synonymous and nonsynonymous substitutions in genes across evolutionary history, biologists can distinguish the evolutionary mechanisms acting on individual genes. Neutral theory provides an explanation for the relatively constant rate of molecular change seen in many species.

- Describe how the ratio of synonymous to nonsynonymous substitutions can be used to determine whether a particular gene region is evolving neutrally, under positive selection, or under stabilizing selection. **See pp. 505–506 and Figure 24.6**
- Contrast two hypotheses for the wide diversity of genome sizes among different organisms. **See pp. 508–509**

We have examined some of the ways that biologists can use genomes to study the molecular mechanisms of evolution. But how do organisms gain new functions through time?

24.3 How Do Genomes Gain and Maintain Functions?

As we noted in the previous section, most multicellular organisms have many more genes than do most unicellular species. But multicellular organisms evolved from unicellular ancestors. Therefore, some mechanisms must exist that result in the increase of gene numbers within genomes over evolutionary time. There are two primary ways to accomplish this increase: genes can be transferred from other species, or genes can be duplicated within species.

Lateral gene transfer can result in the gain of new functions

In Chapter 22, we noted that the tree of life is usually visualized as a branching diagram, with each lineage dividing into two (or more) lineages through time, from one common ancestor to the millions of species that are alive today. Chapter 23 described how, in the process of speciation, ancestral lineages divide into descendant lineages, and it is those speciation events that the tree of life captures. However, there are also processes of **lateral gene transfer**, which allow individual genes, organelles, or fragments of genomes to move horizontally from one lineage to another. Some species may pick up fragments of DNA directly from the environment. Other genes may be picked up in a viral genome and then transferred to a new host when the virus becomes integrated in the new host’s genome. Hybridization between species also results in the transfer of large numbers of genes.

Lateral gene transfer can be highly advantageous to the species that incorporates novel genes from a distant relative. Genes that confer antibiotic resistance, for example, are com-

monly transferred among different species of bacteria. Lateral gene transfer is another way, in addition to mutation and recombination, that species can increase their genetic variability. Genetic variability then provides the raw material on which selection acts, resulting in evolution.

A phylogenetic tree constructed from a single laterally transferred genome fragment is likely to reflect only that transfer event, rather than the overall organismal phylogeny (see Section 26.3). Most biologists prefer to build trees from large samples of genes or their products, so that the underlying species tree (as well as any lateral gene transfer events) can be reconstructed. The depiction of lateral gene transfer events on the underlying species tree are known as *reticulations* on the phylogenetic tree.

The degree to which lateral gene transfer events occur in various parts of the tree of life is a matter of considerable current investigation and debate. Lateral gene transfer appears to be relatively uncommon among most eukaryote lineages, although the two major endosymbioses that gave rise to mitochondria and chloroplasts can be viewed as lateral transfers of entire bacteria genomes to the eukaryote lineage. Some groups of eukaryotes, most notably some plants, are subject to relatively high levels of hybridization among closely related species. Hybridization leads to the exchange of many genes among recently separated lineages of plants. The greatest degree of lateral transfer, however, appears to occur among species of bacteria. Many bacteria genes have been transferred repeatedly among lineages of bacteria, to the point that relationships among the bacteria species are often hard to decipher. Nonetheless, the broad relationships of the major groups of bacteria can still be determined (as we will discuss in Part Seven of this book). Lateral transfer of genes makes it difficult to identify the boundaries of bacteria species, which is one reason why fewer bacteria species have been named than are known to exist.

Most new functions arise following gene duplication

Gene duplication is yet another way in which genomes can acquire new functions. When a gene is duplicated, one copy of that gene is potentially freed from having to perform its original function. The identical copies of a duplicated gene can have any one of four different fates:

- Both copies of the gene may retain their original function (which can result in a change in the amount of gene product that is produced by the organism).
- Both copies of the gene may retain the ability to produce the original gene product, but the expression of the genes may diverge in different tissues or at different times of development.
- One copy of the gene may be incapacitated by the accumulation of deleterious substitutions and become a functionless pseudogene.
- One copy of the gene may retain its original function while the second copy accumulates enough substitutions that it can perform a different function.

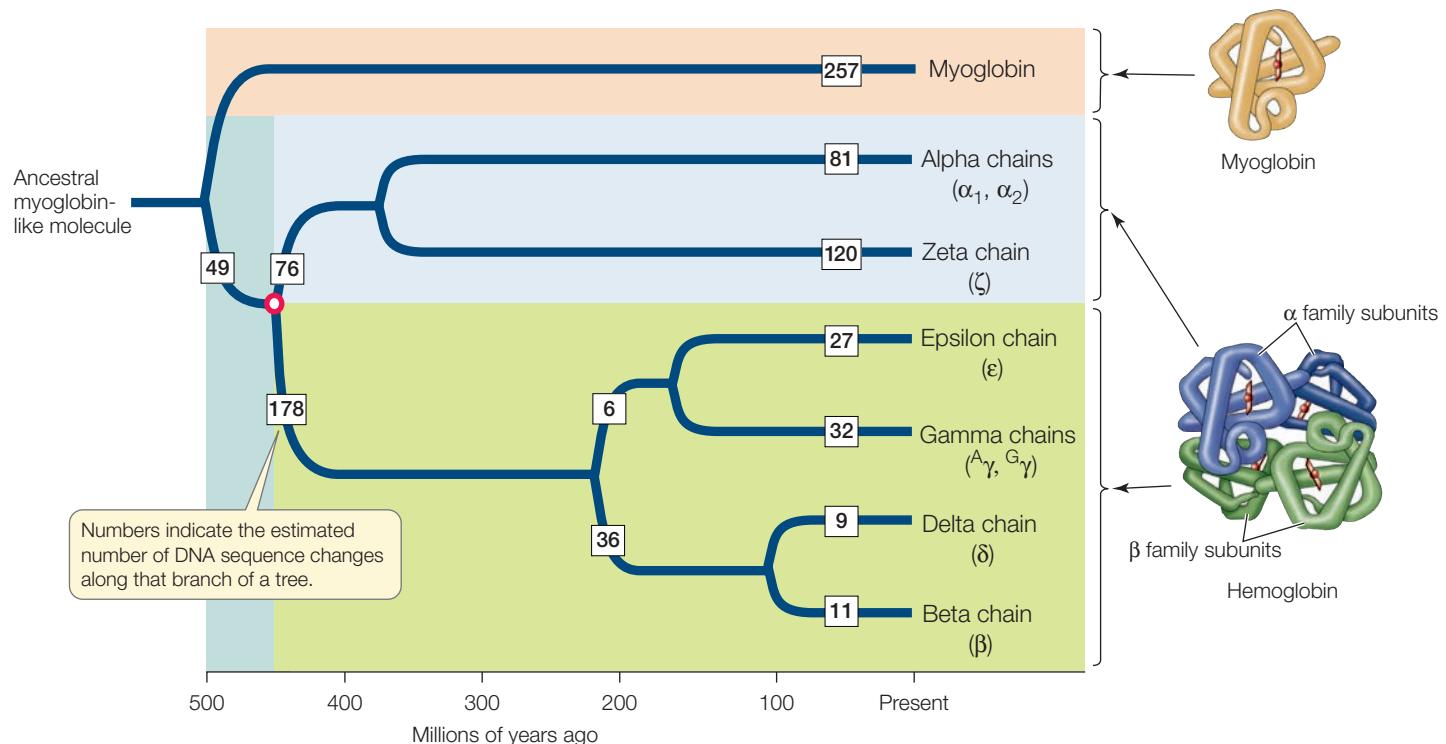
How often do gene duplications arise, and which of these four outcomes is most likely? Investigators have found that rates of gene duplication are fast enough for a yeast or *Drosophila* population to acquire several hundred duplicate genes over the course of a million years. They have also found that most of the duplicated genes in these organisms are very young. Many extra genes are lost from a genome within 10 million years (which is rapid on an evolutionary time scale).

Many gene duplications affect only one or a few genes at a time, but entire genomes are often duplicated in polyploid organisms (including many plants). When all the genes are duplicated, there are massive opportunities for new functions to evolve. That is exactly what appears to have happened in the evolution of vertebrates. The genomes of the jawed vertebrates appear to have four diploid sets of many major genes, which leads biologists to believe that two genome-wide duplication events occurred in the ancestor of these species. These duplications have allowed considerable specialization of individual vertebrate genes, many of which are now highly tissue-specific in their expression. A good example is the duplication of sodium channel genes, which allowed the evolution of the electric organs of electric fishes described in the opening of this chapter.

Although many extra genes disappear rapidly, some duplication events lead to the evolution of genes with new functions. Several successive rounds of duplication and mutation may result in a **gene family**, a group of homologous genes with related functions, often arrayed in tandem along a chromosome. An example of this process is provided by the *globin gene family* (see Figure 17.10). The globins were among the first proteins to be sequenced and compared. Comparisons of their amino acid sequences strongly suggest that the different globins arose via gene duplications. These comparisons can also tell us how long the globins have been evolving separately, because differences among these proteins have accumulated with time.

Hemoglobin, a tetramer (four-subunit molecule) consisting of two α -globin and two β -globin polypeptide chains, carries oxygen in blood. Myoglobin, a monomer, is the primary O_2 storage protein in muscle. Myoglobin's affinity for O_2 is much higher than that of hemoglobin, but hemoglobin has evolved to be more diversified in its role. Hemoglobin binds O_2 in the lungs or gills, where the O_2 concentration is relatively high, transports it to deep body tissues, where the O_2 concentration is low, and releases it in those tissues. With its more complex tetrameric structure, hemoglobin is able to carry four molecules of O_2 , as well as hydrogen ions and carbon dioxide, in the blood.

To estimate the time of the globin gene duplication that gave rise to the α - and β -globin gene clusters, we can create a phylogenetic tree of the gene sequences that encode the various globins (Figure 24.10). The rate of molecular evolution of globin genes has been estimated from other studies, using the divergence times of groups of vertebrates that are well documented in the fossil record. These studies indicate an average rate of divergence for globin genes of about 1 nucleotide substitution every 2 million years. Applying this rate to the gene tree, the two globin gene clusters are estimated to have split about 450 million years ago.



24.10 A Globin Family Gene Tree This gene tree suggests that the α -globin and β -globin gene clusters diverged about 450 million years ago (open circle), soon after the origin of the vertebrates.

yourBioPortal.com

GO TO Web Activity 24.3 • Gene Tree Construction

Some gene families evolve through concerted evolution

Although the members of the globin gene family have diversified in form and function, the members of many other gene families do not evolve independently of one another. For instance, almost all organisms have many copies (up to thousands) of the ribosomal RNA genes. Ribosomal RNA (rRNA) is the principal structural element of ribosomes and, as such, has a primary role in protein synthesis. Every living species needs to synthesize proteins, often in large amounts (especially during early development). Having many copies of the rRNA genes ensures that organisms can rapidly produce many ribosomes and thereby maintain a high rate of protein synthesis.

Like all portions of the genome, ribosomal RNA genes evolve, and differences accumulate in the rRNA genes of different species. But within any one species, the multiple copies of rRNA genes are very similar, both structurally and functionally. This similarity makes sense, because ideally every ribosome in a species should synthesize proteins in the same way. In other words, within a given species, the multiple copies of these rRNA genes are evolving in concert with one another, a phenomenon called **concerted evolution**.

How does concerted evolution occur? There must be one or more mechanisms to cause a substitution in one copy to spread to other copies in a species so that all of the copies remain similar. In fact, two different mechanisms appear to be responsi-

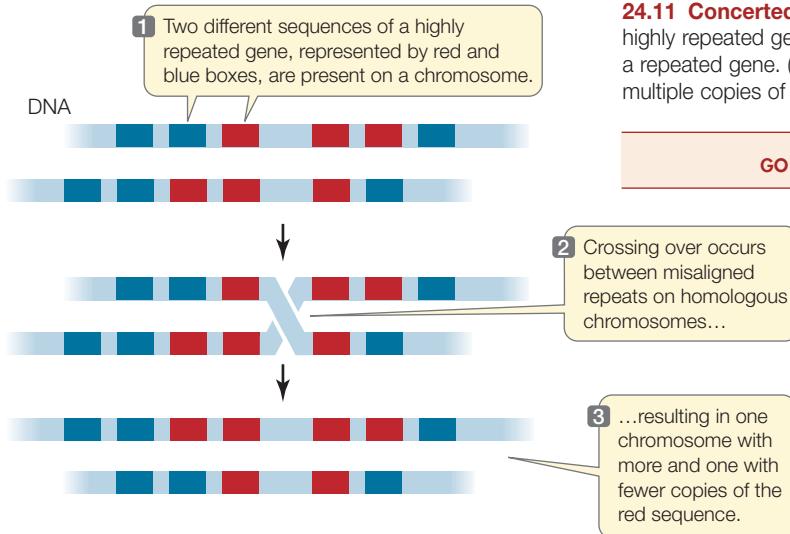
ble for concerted evolution. The first of these is *unequal crossing over*. When DNA is replicated during meiosis in a diploid species, the homologous chromosome pairs align and recombine by crossing over (see Section 11.4).

In the case of highly repeated genes, however, it is easy for genes to become displaced in alignment, since so many copies of the same genes are present in the repeats (Figure 24.11A). The end result is that one chromosome will gain extra copies of the repeat and the other chromosome will have fewer copies of the repeat. If a new substitution arises in one copy of the repeat, it can spread to new copies (or be eliminated) through unequal crossing over. Thus, over time, a novel substitution will either become fixed or lost entirely from the repeat. In either case, all the copies of the repeat will remain very similar to one another.

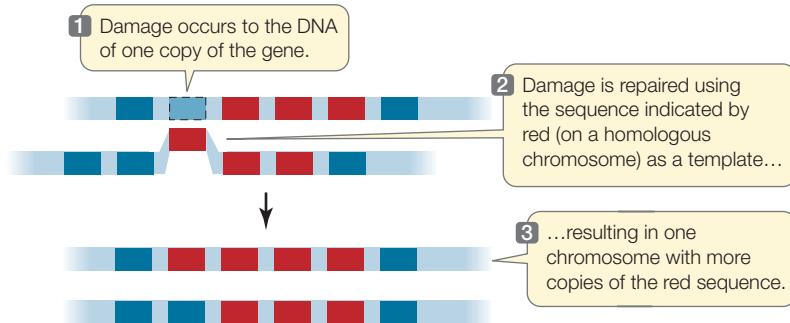
The second mechanism that produces concerted evolution is *biased gene conversion*. This mechanism can be much faster than unequal crossing over, and has been shown to be the primary mechanism for concerted evolution of rRNA genes. DNA strands break often, and are repaired by the DNA repair systems of cells (see Section 13.4). At many times during the cell cycle, the genes for ribosomal RNA are clustered close together. If damage occurs to one of the genes, a copy of the rRNA gene on another chromosome may be used to repair the damaged copy, and the sequence that is used as a template can thereby replace the original sequence (Figure 24.11B). In many cases, this repair system appears to be biased in favor of using particular sequences as templates for repair, and thus the favored sequence rapidly spreads across all copies of the gene. In this way, changes may appear in a single copy and then rapidly spread to all the other copies.

Regardless of the mechanism responsible, the net result of concerted evolution is that the copies of a highly repeated gene do not evolve independently of one another. Mutations still oc-

(A) Unequal crossing over



(B) Biased gene conversion



cur, but once they arise in one copy, they either spread rapidly across all the copies or are lost from the genome completely. This process allows the products of each copy to remain similar through time in both sequence and function.

24.3 RECAP

Gene duplication can lead to the evolution of new functions. Lateral gene transfer can result in the spread of genetic functions between distantly related species. Some highly repeated genes evolve in concert, which maintains uniform functionality.

- Explain the potential advantages of lateral gene transfer. **See pp. 509–510**
- What are four possible outcomes of gene duplication? **See p. 510**
- Describe the pattern of concerted evolution among highly repeated genes and the mechanisms that lead to concerted evolution. **See p. 511 and Figure 24.11**

We have seen how the principles and methods of molecular evolution have opened new vistas in evolutionary science. Now let's consider some of the practical applications of this field.

24.11 Concerted Evolution Two mechanisms can produce concerted evolution of highly repeated genes. (A) Unequal crossing over results in deletions and duplications of a repeated gene. (B) Biased gene conversion can rapidly spread a new variant across multiple copies of a repeated gene.

yourBioPortal.com

GO TO Animated Tutorial 24.1 • Concerted Evolution

24.4 What Are Some Applications of Molecular Evolution?

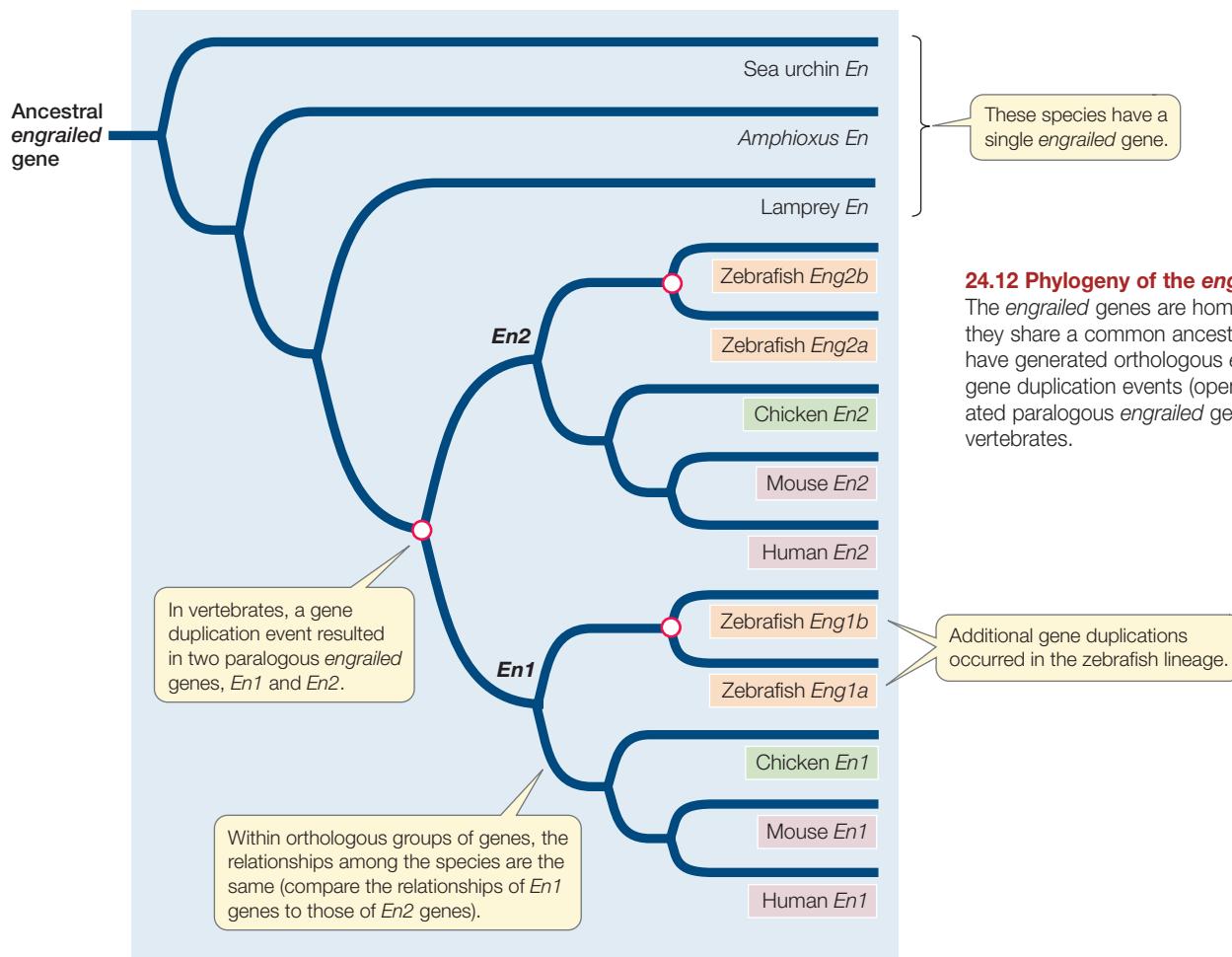
Our understanding of molecular biology has helped reveal how biological molecules function as well as how they diversify. Such knowledge allows scientists to create new molecules with novel functions in the laboratory, and to understand and treat disease.

Molecular sequence data are used to determine the evolutionary history of genes

A **gene tree** shows the evolutionary relationships of a single gene in different species or of the members of a gene family (as in Figure 24.10). The methods for constructing a gene tree are the same as those we described in Section 22.2 for building phylogenetic trees of species. The process involves identifying differences between genes and using those differences to reconstruct the evolutionary history of the genes. Gene trees are often used to infer phylogenetic trees of species, but the two types of trees are not necessarily equivalent. Processes such as gene duplication can give rise to differences between the phylogenetic trees of genes and species. From a gene tree, biologists can reconstruct the history and timing of gene duplication events and learn how gene diversification has resulted in the evolution of new protein functions.

All of the genes of a particular gene family have similar sequences because they have a common ancestry. As we discussed in Section 22.1, features that are similar as a result of common ancestry are referred to as *homologs* of one another. When discussing gene trees, however, we usually need to distinguish between two forms of homology. Genes found in different species, and whose divergence we can trace to the speciation events that gave rise to various species, are called **orthologs**. Genes in the same or different species that are related through gene duplication events are called **paralogs**. When we examine a gene tree, the questions we wish to address determine whether we should compare orthologous or paralogous genes. If we wish to reconstruct the evolutionary history of the species that contain the genes, then our comparison should be restricted to orthologs (because they will reflect the history of speciation events). If we are interested in the changes in function that have resulted from gene duplication events, however, then the appropriate comparison is among paralogs (because they will reflect the history of gene duplication events). If our focus is on the diversification of a gene family through both processes, then we will want to include both paralogs and orthologs in our analysis.

Figure 24.12 depicts a gene tree for the members of a gene family called *engrailed* (its members encode transcription factors



that regulate development). At least three gene duplications have occurred in this family, resulting in up to four different *engrailed* genes (*En*) in some vertebrate species (such as the zebrafish). All of the *engrailed* genes are homologs because they have a common ancestor. Gene duplication events have generated paralogous *engrailed* genes in some lineages of vertebrates. We could compare the orthologous sequences of the *En1* group of genes to reconstruct the history of the bony vertebrates (i.e., all the species in Figure 24.12 except the lamprey), or we could use the orthologous sequences of the *En2* group of genes and expect the same answer (because there is only one history of the underlying speciation events). All bony vertebrates have both groups of *engrailed* genes because the two groups arose from a gene duplication event in the common ancestor of bony vertebrates. If we wanted to focus on the diversification that occurred as a result of this duplication, then the appropriate comparison would be between the paralogous genes of the *En1* versus *En2* groups.

Gene evolution is used to study protein function

Earlier in this chapter we discussed the ways in which biologists can detect regions of genes that are under positive selection for change. What are the practical uses of this information? Consider the evolution of the family of gated sodium channel genes, which we introduced in the opening of this chapter. Sodium channels have many functions, including the control of nerve

impulses in the nervous system. Sodium channels can become blocked by various toxins, such as the tetrodotoxin that is present in puffer fishes and many other animals. A human who eats the tissues of a puffer fish that contain tetrodotoxin can become paralyzed and die, because the tetrodotoxin blocks sodium channels and prevents nerves and muscles from functioning.

Puffer fish themselves have sodium channels; so why doesn't the tetrodotoxin cause paralysis in a puffer fish? The sodium channels of puffer fish (and other animals that sequester tetrodotoxin) have evolved to become resistant to the toxin. Nucleotide substitutions in the puffer fish genome have resulted in changes to the proteins that make up sodium channels, and those changes prevent tetrodotoxin from binding to the sodium channel pore and blocking it. Several different substitutions that result in tetrodotoxin resistance have evolved in the various duplicated sodium channel genes of the many species of puffer fish.

Many other changes that have nothing to do with the evolution of tetrodotoxin resistance have occurred in these genes as well. Biologists who study the function of sodium channels can learn a great deal about how the channels work (and about neurological diseases that are caused by mutations in the sodium channel genes) by understanding which changes have been selected for tetrodotoxin resistance. They can do this by comparing the rates of synonymous and nonsynonymous substitutions across the genes in various lineages that have evolved tetrodo-

24.12 Phylogeny of the *engrailed* Genes

The *engrailed* genes are homologous because they share a common ancestor. Speciation events have generated orthologous *engrailed* genes, and gene duplication events (open circles) have generated paralogous *engrailed* genes among bony vertebrates.

toxin resistance. In a similar manner, molecular evolutionary principles are used to understand function and diversification of function in many other proteins.

As biologists studied the relationship between selection, evolution, and function in macromolecules, they realized that molecular evolution could be used in a controlled laboratory environment to produce new molecules with novel and useful functions. Thus were born the applications of *in vitro* evolution.

In vitro evolution produces new molecules

Living organisms produce thousands of compounds that humans have found useful. The search for such naturally occurring compounds, which can be used for pharmaceutical, agricultural, or industrial purposes, has been termed *bioprospecting*. These compounds are the result of millions of years of molecular evolution across millions of species of living organisms. Yet biologists can imagine molecules that could have evolved but have not, lacking the right combination of selection pressures and opportunities.

For instance, we might like to have a molecule that binds a particular environmental contaminant so that it can be easily isolated and extracted from the environment. But if the environmental contaminant is synthetic (not produced naturally), then it is unlikely that any living organism would have evolved a molecule with the function we desire. This problem was the inspiration for the field of ***in vitro* evolution**, in which new molecules are produced in the laboratory to perform novel and useful functions.

The principles of *in vitro* evolution are based on the principles of molecular evolution that we have learned from the natural world. Consider the evolution of a new RNA molecule that was produced in the laboratory using the principles of mutation and selection. This molecule's intended function was to join two other RNA molecules (acting as a ribozyme with a function similar to that of the naturally occurring DNA ligase described in Section 13.3, but for RNA molecules). The process started with a large pool of random RNA sequences (10^{15} different sequences, each about 300 nucleotides long), which were then selected for any ligase activity (Figure 24.13). None were very effective ribozymes for ligase activity, but some were slightly better than others. The best of the ribozymes were selected and reverse-transcribed into cDNA (using the enzyme reverse transcriptase). The cDNA molecules were then amplified using the polymerase chain reaction (PCR; see Figure 13.22).

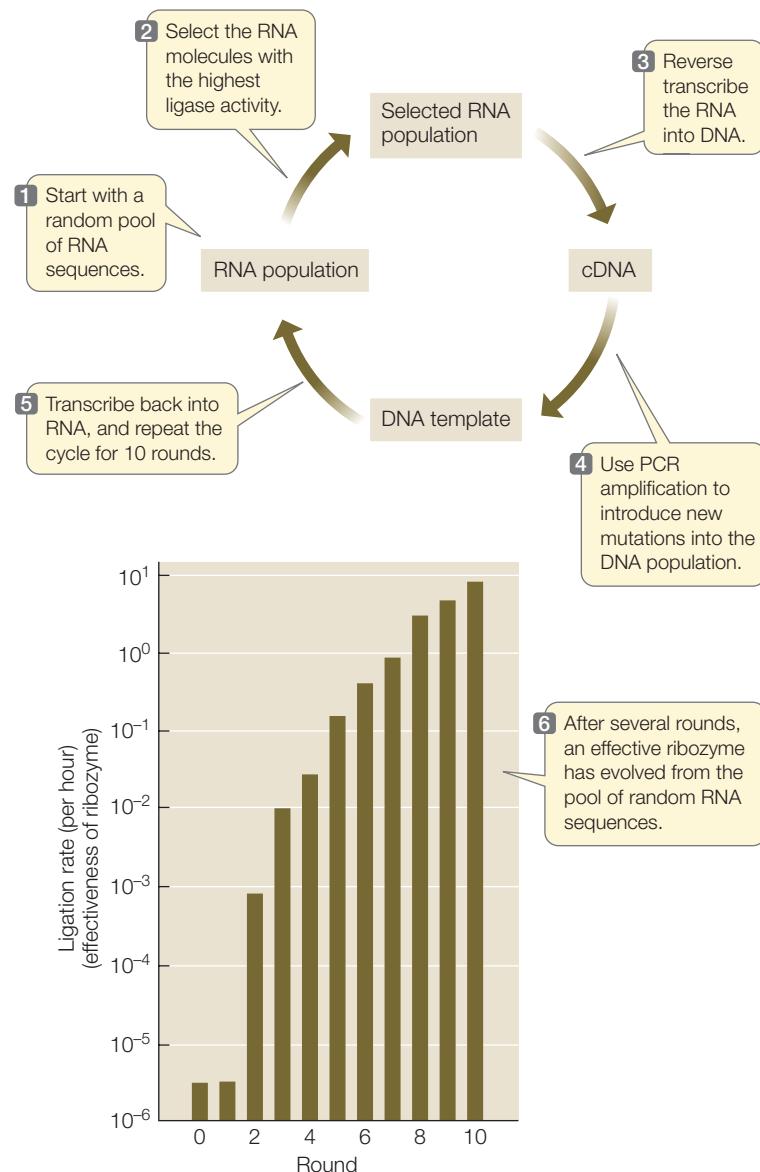
PCR amplification is not perfect, and it introduced many new mutations into the pool of sequences. These sequences were then transcribed back into RNA molecules using RNA polymerase, and the process was repeated. The ligase activity of the RNAs evolved quickly; after 10 rounds of *in vitro* evolution, it had increased by about 7 million times (see Figure 24.13). Similar techniques have since been used to create a wide variety of molecules with novel enzymatic and binding functions.

24.13 In Vitro Evolution Starting with a large pool of random RNA sequences, Bartel and Szostak produced a new ribozyme through rounds of mutation and selection for the ability to ligate RNA sequences.

Molecular evolution is used to study and combat diseases

Many of the most problematic human diseases are caused by living, evolving organisms that present a moving target for modern medicine. Recall the example of influenza described at the start of Chapter 21 and that of HIV described in Chapter 22. The control of these and many other human diseases depends on techniques that can track the evolution of pathogenic organisms through time.

During the past century, transportation advances have allowed humans to move around the world with unprecedented speed and increasing frequency. Unfortunately, this mobility has allowed pathogens to be transmitted among human populations at much higher rates, which has led to the global emergence of many “new” diseases. Most of these emerging diseases are caused by viruses, and virtually all new viral diseases have been identified by evolutionary comparison of their genomes with those of known viruses. In recent years, for example, ro-



dent-borne hantaviruses have been identified as the source of widespread respiratory illnesses, and the virus (and its host) that causes Sudden Acute Respiratory Syndrome (SARS) has been identified using evolutionary comparisons of genes. Studies of the origins, the timing of emergence, and the global diversity of many human pathogens depend on the principles of molecular evolution, as do the efforts to develop and use effective vaccines against these pathogens. For example, the techniques to develop polio vaccines, as well as the methods used to track their effectiveness in human populations, rely on molecular evolutionary approaches.

In the future, molecular evolution will become even more critical to the identification of human (and other) diseases. Once biologists have collected data on the genomes of enough organisms, it will be possible to identify an infection by sequencing a portion of the infecting organism's genome and comparing this sequence with other sequences on an evolutionary tree. At present, it is difficult to identify many common viral infections (those that cause "colds," for instance). As genomic databases and evolutionary trees increase, however, automated methods of sequencing and rapid phylogenetic comparison of the sequences will allow us to identify and treat a much wider array of human illnesses.

24.4 RECAP

Molecular evolutionary studies have provided biologists with new tools to understand the functions of macromolecules and how those functions can change through time. Molecular evolution is used to develop synthetic molecules for industrial and pharmaceutical uses and to identify and combat human diseases.

- Why might a biologist limit a particular investigation to orthologous (as opposed to paralogous) genes? **See pp. 512–513**
- Explain how gene evolution can be used to study protein function. **See pp. 513–514**
- Describe the process of in vitro evolution. **See p. 514 and Figure 24.13**

Now that we have discussed how organisms and biological molecules evolve, we are ready to consider the evolutionary history of the Earth. Chapter 25 describes the long-term evolutionary changes that have given rise to all of life's diversity.

CHAPTER SUMMARY

24.1 How Are Genomes Used to Study Evolution?

SEE WEB ACTIVITY 24.1

- The field of **molecular evolution** concerns relationships between the structures of genes and proteins and the functions of organisms.
- A **genome** is an organism's full set of genes and noncoding DNA. In eukaryotes, the genome includes genetic material in the nucleus of the cell as well as in mitochondria and chloroplasts (where present).
- Nucleotide substitutions may or may not result in amino acid replacements in the encoded proteins.
- The estimated number of substitutions between sequences can be calculated from a **similarity matrix** using models of sequence evolution that account for changes that cannot be observed directly. **Review Figure 24.1, WEB ACTIVITY 24.2**
- The concept of homology (similarity that results from common ancestry) extends down to the level of particular positions in nucleotide or amino acid sequences. **Sequence alignments** from different organisms allow us to compare the sequences and identify homologous positions. **Review Figure 24.3**

24.2 What Do Genomes Reveal about Evolutionary Processes?

- **Nonsynonymous substitutions** of nucleotides result in amino acid replacements in proteins, but **synonymous substitutions** do not. **Review Figure 24.5**
- Rates of synonymous substitution are typically higher than rates of nonsynonymous substitution in protein-coding genes (a result of stabilizing selection). **Review Figure 24.6**

- Much of the molecular change in nucleotide sequences is a result of **neutral evolution**. The rate of fixation of neutral mutations is independent of population size and is equal to the mutation rate.
- Positive selection for change in a protein-coding gene may be detected by a higher rate of nonsynonymous versus synonymous substitutions.
- Genome size evolves by the addition or deletion of genes and noncoding DNA. The total size of genomes varies much more widely across multicellular species than does the number of functional genes. **Review Figures 24.8 and 24.9**
- Even though many noncoding regions of the genome may not have direct functions, these regions can affect the phenotype of an organism by influencing gene expression.
- Functionless **pseudogenes** can serve as the raw material for the evolution of new genes.

24.3 How Do Genomes Gain and Maintain Functions?

- Lateral gene transfer can result in the rapid acquisition of new functions from distantly related species.
- **Gene duplications** can result in increased production of the gene's product, in pseudogenes, or in new gene functions. Several rounds of gene duplication can give rise to multiple genes with related functions, known as a **gene family**.
- Some highly repeated genes evolve by **concerted evolution**: multiple copies within an organism maintain high similarity, while the genes continue to diverge between species. **SEE ANIMATED TUTORIAL 24.1**

24.4 What Are Some Applications of Molecular Evolution?

SEE WEB ACTIVITY 24.3

- **Gene trees** describe the evolutionary history of particular genes or gene families.
- **Orthologs** are genes that are related through speciation events, whereas **paralogs** are genes that are related through gene duplication events. **Review Figure 24.12**

- Protein function can be studied by examining gene evolution. Detection of positive selection can be used to identify molecular changes that have resulted in functional changes.
- **In vitro evolution** is used to produce synthetic molecules with particular desired functions. **Review Figure 24.13**
- Many diseases are identified, studied, and combated through molecular evolutionary investigations.

SELF-QUIZ

1. A higher rate of synonymous relative to nonsynonymous substitutions in a protein-coding gene is expected under
 - a. purifying selection.
 - b. positive selection.
 - c. neutral evolution.
 - d. concerted evolution.
 - e. none of the above
2. Before nucleotide and amino acid sequences can be compared in an evolutionary framework, they must be aligned to account for
 - a. deletions and insertions.
 - b. selection and neutrality.
 - c. parallelisms and convergences.
 - d. gene families.
 - e. all of the above
3. Models of nucleotide sequence evolution, developed by biologists to estimate sequence divergence, include parameters that account for
 - a. substitution rates between different nucleotides.
 - b. differences in substitution rates across different positions in a gene.
 - c. differences in nucleotide frequencies.
 - d. all of the above
 - e. none of the above
4. The rate of fixation of neutral mutations is
 - a. independent of population size.
 - b. higher in small populations than in large populations.
 - c. higher in large populations than in small populations.
 - d. slower than the rate of fixation of deleterious mutations.
 - e. none of the above
5. Genome size differs widely among different multicellular species of eukaryotes. What is the greatest contributing cause for these differences?
 - a. The number of protein-coding genes
 - b. The amount of noncoding DNA
 - c. The number of duplicated genes
 - d. The degree of concerted evolution
 - e. The amount of positive selection for change in protein-coding genes
6. Which of the following is *not* true of concerted evolution?
 - a. Concerted evolution refers to the nonindependent evolution of some repeated genes within a species.
 - b. Unequal crossing over may produce concerted evolution.
 - c. Biased gene conversion may produce concerted evolution.
 - d. Ribosomal RNA genes are an example of a gene family that has undergone concerted evolution.
 - e. Concerted evolution results in divergence of members of a gene family within an organism.
7. When a gene is duplicated, which of the following may occur?
 - a. Production of the gene's product may increase.
 - b. The two copies may become expressed in different tissues.
 - c. One copy of the gene may accumulate deleterious substitutions and become functionless.
 - d. The two copies may diverge and acquire different functions.
 - e. All of the above
8. Paralogous genes are genes that trace back to a common
 - a. speciation event.
 - b. substitution event.
 - c. insertion event.
 - d. deletion event.
 - e. duplication event.
9. Which of the following is true of *in vitro* evolution?
 - a. *In vitro* evolution refers to bioprospecting for naturally occurring macromolecules.
 - b. *In vitro* evolution can produce new molecular sequences not known from nature.
 - c. *In vitro* evolution can only produce new proteins.
 - d. *In vitro* evolution only selects for changes that were present in the starting pool of molecules, and does not introduce any new mutations.
 - e. All of the above
10. Which of the following is true of the use of molecular evolutionary studies of human disease?
 - a. Molecular evolutionary studies are useful for identifying many diseases.
 - b. Molecular evolutionary studies are often used to determine the origin of emerging diseases.
 - c. Molecular evolutionary studies are important for developing vaccines against diseases.
 - d. Molecular evolutionary studies are used to track the effectiveness of polio vaccines in human populations.
 - e. All of the above

FOR DISCUSSION

1. Rates of evolutionary change differ among different molecules, and different species differ widely in generation times and population sizes. How does this variation limit how and in what ways we can use the concept of a molecular clock to help us answer questions about the evolution of both molecules and organisms?
2. One hypothesis proposed to explain the existence of large amounts of noncoding DNA is that the cost of maintaining that DNA is so small that natural selection is too weak to reduce it. How could you test this hypothesis against the hypothesis that genome size is functionally related to developmental rate?
3. If fossil evidence and molecular evidence disagree on the date of a major lineage split, which of the two kinds of evidence would you favor? Why?
4. Scientists can produce and release into the wild genetically modified mosquitoes that are unable to harbor and transmit malarial parasites. What ethical issues need to be discussed before such releases are permitted?

ADDITIONAL INVESTIGATION

Over evolutionary history, many groups of organisms that inhabit caves have lost the organs of sight. For instance, although surface-dwelling crayfishes have functional eyes, several crayfish species that are restricted to underground habitats lack eyes. Opsins are a group of light-sensitive proteins known to have an important function in vision (see Chapter 46), and opsin genes are expressed in eye tissues. Opson genes are pres-

ent in the genomes of eyeless, cave-dwelling crayfish. Two alternative hypotheses are (1) the opson genes are no longer experiencing purifying selection (because there is no longer selection for function in vision); or (2) the opson genes are experiencing selection for a function other than vision. How would you investigate these alternatives using the sequences of the opson genes in various species of crayfishes?

When hawk-sized dragonflies ruled the air

Almost anyone who has spent time around fresh water ponds is familiar with dragonflies. Their hovering flight, bright colors, and transparent wings stimulate our visual senses on bright summer afternoons as they fly about their business of devouring mosquitoes, mating, and laying their eggs. The largest dragonflies alive today have wingspans that can be covered by a human hand. Three hundred million years ago, however, dragonflies such as *Meganeuropsis permiana* had wingspans of more than 70 centimeters—well over 2 feet, matching or exceeding the wingspans of many modern birds of prey—and were the largest flying predators on Earth.

No flying insects alive today are anywhere near this size. But during the Carboniferous and Permian geological periods, between 350 and 250 million years ago, many groups of flying insects contained gigantic members. *Meganeuropsis* probably ate huge mayflies and other

giant flying insects that shared their home in the Permian swamps. These enormous insects were themselves eaten by giant amphibians.

None of the giant flying insects or amphibians of that time would be able to survive on Earth today. The oxygen concentrations in Earth's atmosphere were about 50 percent higher than they are now, and those high oxygen levels are thought to have been necessary to support giant insects and their huge amphibian predators.

Paleontologists have uncovered fossils of *Meganeuropsis permiana* in the rocks of Kansas. How do we know the age of these fossils, and how can we know how much oxygen that long-vanished atmosphere contained? The *stratigraphic* layering of the rocks allows us to tell their ages relative to each other, but it does not by itself indicate a given layer's absolute age.

One of the remarkable achievements of twentieth-century scientists was to develop sophisticated techniques that use the decay rates of various radio-isotopes, changes in Earth's magnetic field, and the ratios of certain molecules to infer conditions and events in the remote past and to date them accurately. It is those methods that allow us to age the fossils of *Meganeuropsis* and to calculate the concentration of oxygen in Earth's atmosphere at the time.

The development of the science of biology is intimately linked to changing concepts of



Giant Dragonflies *Meganeuropsis permiana*, shown here in a reconstruction from fossils, dwarfed modern dragonflies (shown in the inset at the same scale) in size. Otherwise, however, the Permian giant was quite similar to modern dragonflies in general appearance.



Younger Rocks Lie on Top of Older Rocks In the Grand Canyon, the Colorado River cut through and exposed many strata of ancient rocks. The oldest rocks visible here formed about 540 million years ago. The youngest, at the top, are about 500 million years old. Knowing the ages of rock strata allows scientists to date the fossils found in each stratum.

time, especially of the age of Earth. About 150 years ago, geologists first provided solid evidence that Earth is ancient; before 1850, most people believed it was no more than a few thousand years old. For many more years, physicists continued to underestimate Earth's age, until an understanding of radioactive decay was developed. Today we know that Earth is about 4.5 billion years old and that life has existed on it for about 3.8 billion of those years. That means human civilizations have occupied Earth for less than 0.0003 percent of the history of life. Discovering what happened before humans were around is an ongoing and exciting area of science.

IN THIS CHAPTER we will examine how biologists assign dates to events in the distant evolutionary past, and how such dating allows us to review the major changes in physical conditions on Earth during the past 4 billion years. We will then look at how these changes in physical conditions have influenced the major patterns in the evolution of life, and describe how scientists organize our knowledge of biological diversity based on the relationships among species.

CHAPTER OUTLINE

- 25.1 How Do Scientists Date Ancient Events?**
- 25.2 How Have Earth's Continents and Climates Changed over Time?**
- 25.3 What Are the Major Events in Life's History?**

25.1 How Do Scientists Date Ancient Events?

Many evolutionary changes happen rapidly enough to be studied directly and manipulated experimentally. Plant and animal breeding by agriculturalists and insects' evolution of resistance to pesticides are examples of rapid, short-term evolution. Other changes, such as the appearance of new species and evolutionary lineages, usually take place over much longer time frames.

To understand the long-term patterns of evolutionary change, we must think in time frames spanning many millions of years, and consider events and conditions very different from those we observe today. Earth of the distant past was so unlike the present that it seems like a foreign planet inhabited by strange organisms. The continents were not where they are today, and climates were sometimes dramatically different from those of today.

Fossils—the preserved remains of ancient organisms—can tell us a great deal about the body form, or *morphology*, of organisms that lived long ago, as well as how and where they lived. Fossils provide a direct record of evolution. But to understand patterns of evolutionary change, we must also understand how Earth has changed over time.

Earth's history is largely recorded in its rocks. We cannot tell the ages of rocks just by looking at them, but we can determine the ages of rocks relative to one another. The first person to formally recognize that this could be done was the seventeenth-century Danish physician Nicolaus Steno. Steno realized that in undisturbed **sedimentary rocks** (rocks formed by the accumulation of grains on the bottom of bodies of water), the oldest layers of rock, or **strata** (singular *stratum*), lie at the bottom; thus successively higher strata are progressively younger.

Geologists, particularly the eighteenth-century English scientist William Smith, subsequently combined Steno's insight with their observations of fossils contained in sedimentary rocks. They concluded that:

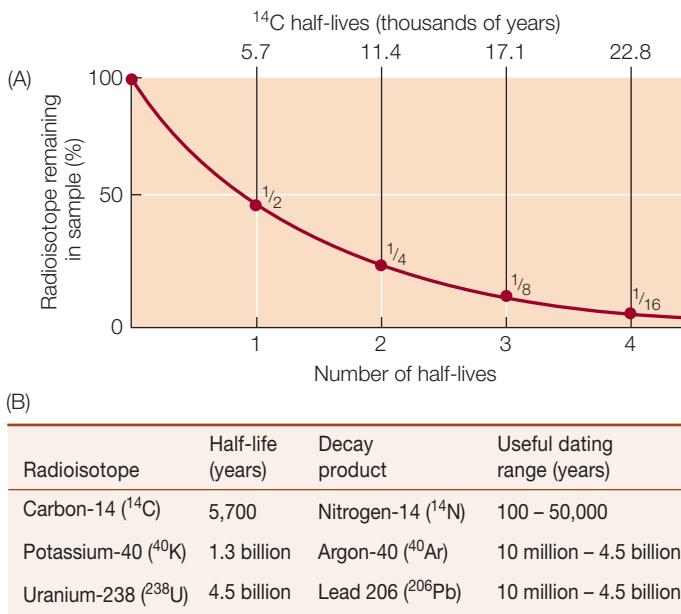
- Fossils of similar organisms are found in widely separated places on Earth.
- Certain fossils are always found in younger rocks, and certain other fossils in older rocks.
- Organisms found in higher, more recent strata are more similar to modern organisms than are those found in lower, more ancient strata.

These patterns revealed much about the relative ages of sedimentary rocks as well as patterns in the evolution of life. But the geologists still could not tell how old the rocks were. A method of dating rocks did not become available until after radioactivity was discovered at the beginning of the twentieth century.

Radioisotopes provide a way to date rocks

Radioactive isotopes of atoms (see Section 2.1) decay in a predictable pattern over long time periods. During each successive time interval, known as a **half-life**, half of the remaining radioactive material of the radioisotope decays to become a different, stable isotope (Figure 25.1A).

To use a radioisotope to date a past event, we must know or estimate the concentration of the isotope at the time of that event. In the case of carbon, the production of new carbon-14 (^{14}C) in the upper atmosphere (by the reaction of neutrons with nitrogen-14) just balances the natural radioactive decay of ^{14}C to ^{14}N . Therefore, the ratio of ^{14}C to its stable isotope, carbon-12 (^{12}C), is relatively constant in living organisms and their environment. As soon as an organism dies, however, it ceases to exchange carbon compounds with its environment. Its decaying ^{14}C is no longer replenished, and the ratio of ^{14}C to ^{12}C in its remains decreases through time. Paleontologists can use the ratio of ^{14}C to ^{12}C in fossil material to date fossils that are less than 50,000 years old (and thus the sedimentary rocks that contain those fossils). If fossils are older than that, so little ^{14}C remains that the limits of detection using this particular isotope are reached.



25.1 Radioactive Isotopes Allow Us to Date Ancient Rocks The decay of radioactive “parent” atoms into stable “daughter” isotopes happens at a steady rate known as a half-life. (A) The graph demonstrates the principle of half-life using carbon-14 (^{14}C) as an example. (B) Radioisotopes have different characteristic half-lives that allow us to measure how much time has elapsed since the rocks containing them were laid down.

TABLE 25.1

Earth's Geological History

RELATIVE TIME SPAN	ERA	PERIOD	ONSET	MAJOR PHYSICAL CHANGES ON EARTH
Precambrian	Cenozoic	Quaternary	2.6 mya	Cold/dry climate; repeated glaciations
		Tertiary	65 mya	Continents near current positions; climate cools
Precambrian	Mesozoic	Cretaceous	145 mya	Northern continents attached; Gondwana begins to drift apart; meteorite strikes Yucatán Peninsula
		Jurassic	200 mya	Two large continents form: Laurasia (north) and Gondwana (south); climate warm
		Triassic	251 mya	Pangaea begins to slowly drift apart; hot/humid climate
Precambrian	Paleozoic	Permian	297 mya	Extensive lowland swamps; O_2 levels 50% higher than present; by end of period continents aggregate to form Pangaea, and O_2 levels begin to drop rapidly
		Carboniferous	359 mya	Climate cools; marked latitudinal climate gradients
		Devonian	416 mya	Continents collide at end of period; meteorite probably strikes Earth
	Precambrian	Silurian	444 mya	Sea levels rise; two large land masses emerge; hot/humid climate
		Ordovician	488 mya	Massive glaciation, sea level drops 50 meters
		Cambrian	542 mya	O_2 levels approach current levels
			900 mya	O_2 level at >5% of current level
			1.5 bya	O_2 level at >1% of current level
			3.8 bya	O_2 first appears in atmosphere
			4.5 bya	

Note: mya, million years ago; bya, billion years ago.

Radioisotope dating methods have been expanded and refined

Sedimentary rocks are formed from materials that existed for varying lengths of time before being transported, sometimes over long distances, to the site of their deposition. Therefore, the inorganic isotopes in a sedimentary rock do not contain reliable information about the date of its formation. Dating rocks more ancient than 50,000 years requires estimating isotope concentrations in *igneous* rocks—rocks formed when molten material cools. To date older sedimentary rocks, geologists search for places where sedimentary rocks show igneous intrusions of volcanic ash or lava flows.

A preliminary estimate of the age of an igneous rock determines which isotope is used to date it (**Figure 25.1B**). The decay of potassium-40 (which has a half-life of 1.3 billion years) to argon-40 has been used to date many of the ancient events in the evolution of life. Fossils in the adjacent sedimentary rock that are similar to those in other rocks of known ages provide additional clues.

Radioisotope dating of rocks, combined with fossil analysis, is the most powerful method of determining geological age. But in places where sedimentary rocks do not contain suitable ig-

MAJOR EVENTS IN THE HISTORY OF LIFE

Humans evolve; many large mammals become extinct
Diversification of birds, mammals, flowering plants, and insects
Dinosaurs continue to diversify; mass extinction at end of period (~76% of species disappear)
Diverse dinosaurs; radiation of ray-finned fishes; first fossils of flowering plants
Early dinosaurs; first mammals; marine invertebrates diversify; mass extinction at end of period (~65% of species disappear)
Reptiles diversify; giant amphibians and flying insects present; mass extinction at end of period (~96% of species disappear)
Extensive “fern” forests; first reptiles; insects diversify
Fishes diversify; first insects and amphibians; mass extinction at end of period (~75% of species disappear)
Jawless fishes diversify; first ray-finned fishes; plants and animals colonize land
Mass extinction at end of period (~75% of species disappear)
Rapid diversification of multicellular animals; diverse photosynthetic protists
Ediacaran fauna; earliest fossils of multicellular animals
Eukaryotes evolve
Origin of life; prokaryotes flourish

neous intrusions and few fossils are present, paleontologists turn to other methods.

One method, known as **paleomagnetic dating**, relates the ages of rocks to patterns in Earth's magnetism, which change over time. Earth's magnetic poles move and occasionally reverse themselves. Because both sedimentary and igneous rocks preserve a record of Earth's magnetic field at the time they were formed, paleomagnetism helps determine the ages of those rocks. Other dating methods use information about continental drift, sea level changes, and molecular clocks (the last of which is described in Section 22.3).

Using these methods, geologists divided the history of life into *eras*, which in turn are subdivided into *periods* (**Table 25.1**). The boundaries between these time frames are based on striking differences scientists have observed in the assemblages of fossil organisms contained in successive layers of rocks. Geologists defined and named these divisions before they were able to establish the ages of fossils, adding and refining the time scales as new methods for geological dating were developed.

25.1 RECAP

Fossils in sedimentary rocks enabled geologists to determine the relative ages of organisms, but absolute dating was not possible until the discovery of radioactivity. Geologists divide the history of life into eras and periods, based on assemblages of fossil organisms found in successive layers of rocks.

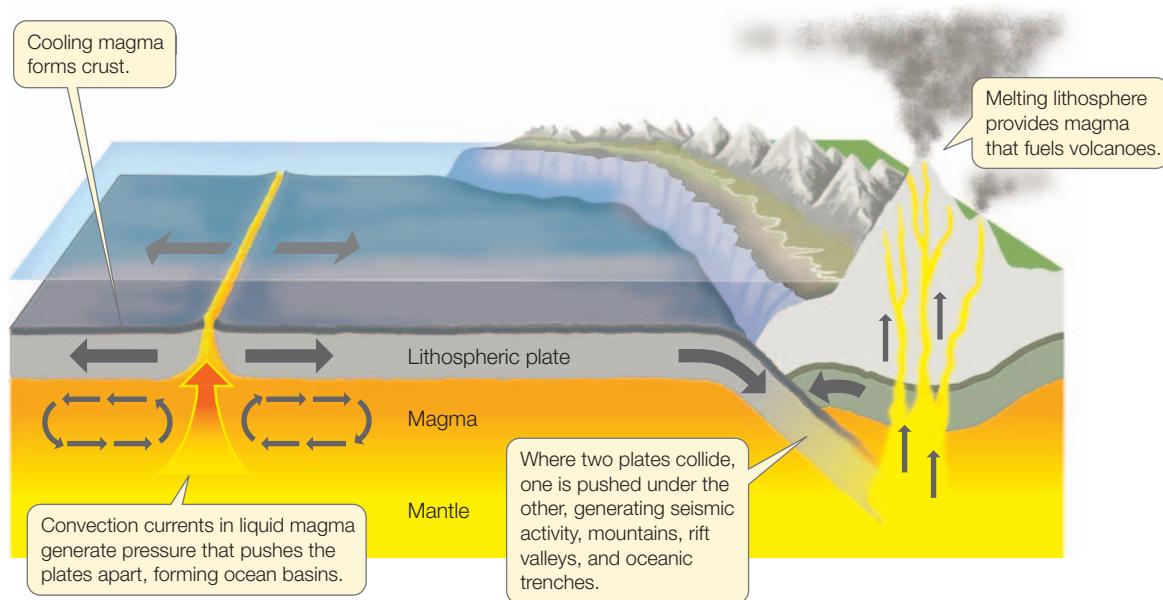
- What observations about fossils suggested to geologists that they could be used to determine the relative ages of rocks? **See p. 519**
- How is the rate of decay of radioisotopes used to estimate the absolute ages of rocks? **See p. 520 and Figure 25.1**

The scale at the left of Table 25.1 gives a relative sense of geological time, especially the vast expanse of the Precambrian era, during which early life evolved amid stupendous physical changes of Earth and its atmosphere. During the Precambrian to Cambrian transition, an “explosion” of new life forms took place as representatives of many of the major multicellular groups of life evolved. Earth continued to undergo massive physical changes that influenced the evolution of life, and these events and important milestones are listed in the table. In the next two sections we'll discuss the most important of these changes.

25.2 How Have Earth's Continents and Climates Changed over Time?

The globes and maps that adorn our walls, shelves, and books give an impression of a static Earth. It would be easy for us to assume that the continents have always been where they are, but we would be wrong. The idea that Earth's land masses have changed position over the millennia, and that they continue to do so, was first put forth in 1912 by the German meteorologist

25.2 Plate Tectonics and Continental Drift The heat of Earth's core generates convection currents (arrows) in the magma that push the lithospheric plates, along with the land masses lying on them, together or apart. When lithospheric plates collide, one often slides under the other. The resulting seismic activity can create mountains and deep rift valleys (the latter known as trenches when they occur under ocean basins).



and geophysicist Alfred Wegener. His book *The Origin of Continents and Oceans* was initially met with skepticism and resistance. By the 1960s, however, physical evidence and increased understanding of the geophysics of **plate tectonics**—the study of movement of major land masses—had convinced virtually all geologists of the reality of Wegener's vision.

Earth's crust consists of several solid plates approximately 40 kilometers thick, which collectively make up the lithosphere. The lithospheric plates float on a fluid layer of molten rock, or magma (Figure 25.2). Heat produced by radioactive decay deep in Earth's core sets up convection currents in the fluid magma, which then rises and exerts tremendous pressure on the solid plates. When the pressure of the rising magma pushes plates apart, ocean basins may form between them. When plates are pushed together, they either move sideways past each other or one plate slides under the other, pushing up mountain ranges and carving deep rift valleys. When they occur under the water of ocean basins, rift valleys are known as trenches. The

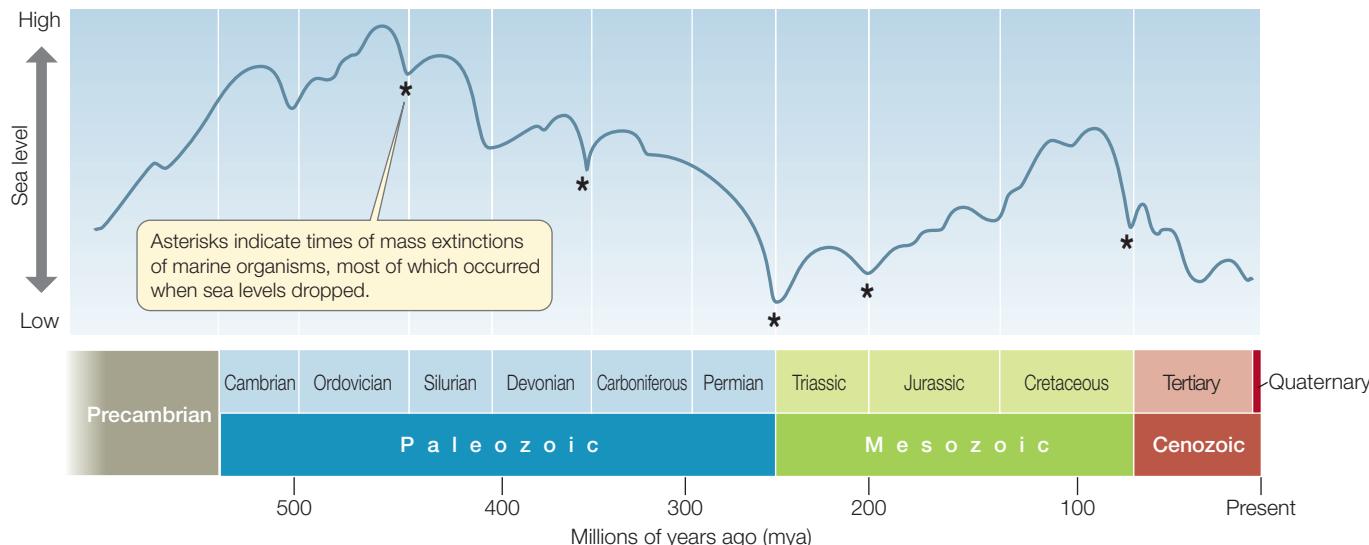
movement of the lithospheric plates and the continents they contain is known as **continental drift**.

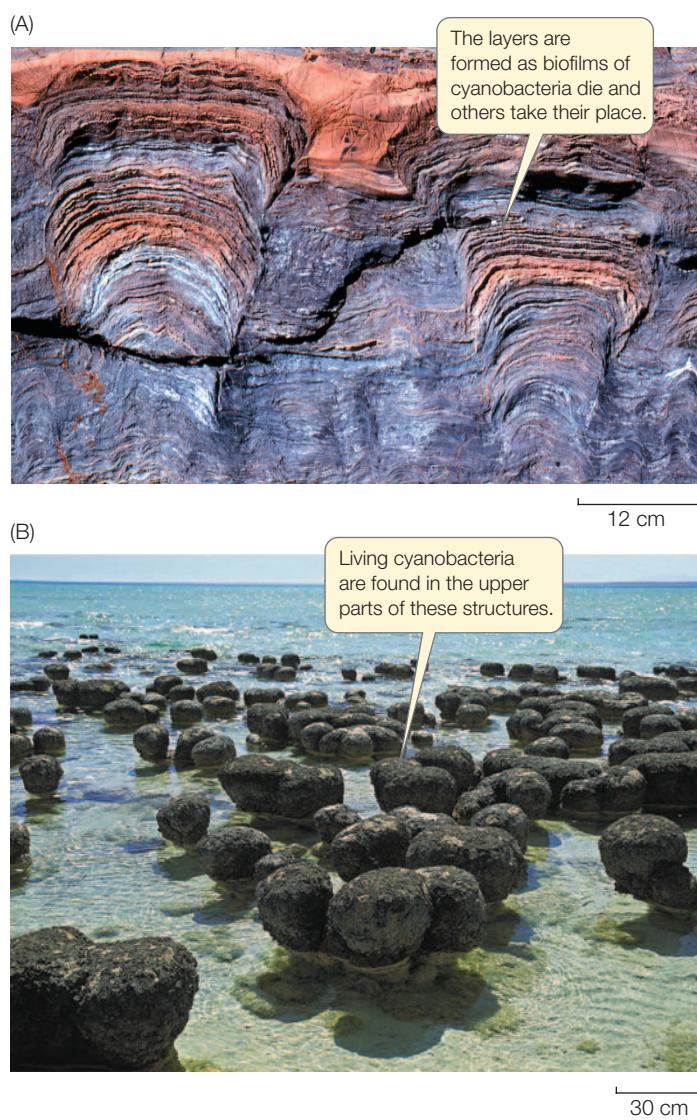
We now know that at times the drifting of the plates has brought continents together and at other times has pushed them apart (these movements are depicted in Figure 25.12). The positions and sizes of the continents influence oceanic circulation patterns, global climates, and sea levels. Major drops in sea level have usually been accompanied by massive extinctions—particularly of marine organisms, which could not survive the exposure of vast areas of the continental shelves and the disappearance of the shallow seas that covered them (Figure 25.3).

yourBioPortal.com

GO TO Animated Tutorial 25.1 • Evolution of the Continents

25.3 Sea Levels Have Changed Repeatedly Most mass extinctions of marine organisms (indicated by asterisks) have coincided with periods of low sea levels.





25.4 Stromatolites (A) A vertical section through a fossil stromatolite. (B) These rocklike structures are living stromatolites that thrive in the very salty waters of Shark Bay in western Australia. Layers of cyanobacteria are found in the uppermost parts of the structures.

Oxygen concentrations in Earth's atmosphere have changed over time

As the continents have moved over Earth's surface, the world has experienced other physical changes, including large increases and decreases in atmospheric oxygen. The atmosphere of early Earth probably contained little or no free oxygen gas (O_2). The increase in atmospheric O_2 came in two big steps more than a billion years apart. The first step occurred at least 2.4 billion years ago (bya), when certain bacteria evolved the ability to use water as the source of hydrogen ions for photosynthesis. By chemically splitting H_2O , these bacteria generated atmospheric O_2 as a waste product. They also made electrons available for reducing CO_2 to form organic compounds (see Section 10.3).

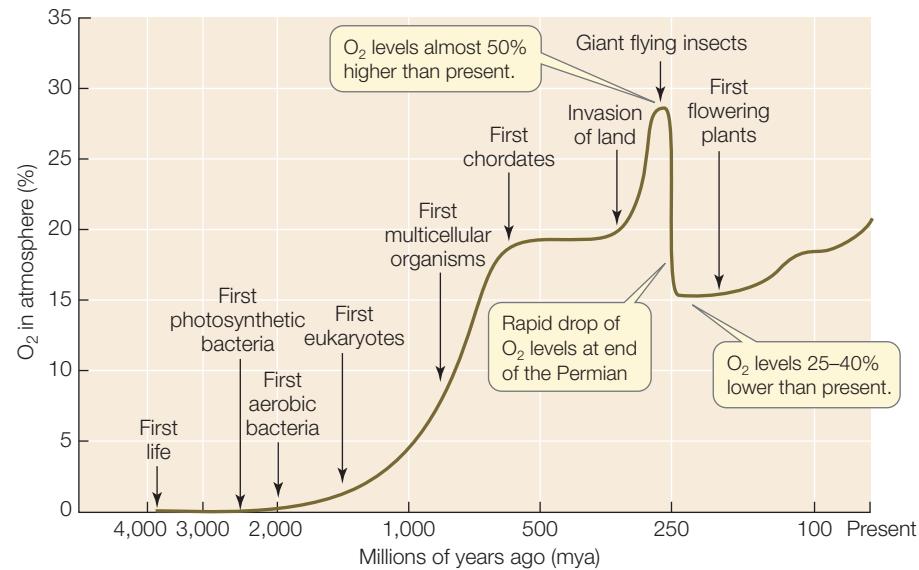
One group of O_2 -generating bacteria, the *cyanobacteria*, formed rocklike structures called

stromatolites, which are abundantly preserved in the fossil record. Cyanobacteria are still forming stromatolites today in a few very salty places on Earth (Figure 25.4). Cyanobacteria liberated enough O_2 to open the way for the evolution of oxidation reactions as the energy source for the synthesis of ATP (see Section 9.1).

The evolution of life thus irrevocably changed the physical nature of Earth. Those physical changes, in turn, influenced the evolution of life. When it first appeared in the atmosphere, O_2 was poisonous to the anaerobic prokaryotes that inhabited Earth at the time. Over millennia, however, prokaryotes that evolved the ability to metabolize O_2 not only survived but gained several advantages. Aerobic metabolism proceeds more rapidly and harvests energy more efficiently than anaerobic metabolism (see Section 9.4), and organisms with aerobic metabolism replaced anaerobes in most of Earth's environments.

An atmosphere rich in O_2 also made possible larger cells and more complex organisms. Small unicellular aquatic organisms can obtain enough O_2 by simple diffusion even when O_2 concentrations are very low. Larger unicellular organisms have lower surface area-to-volume ratios (see Figure 5.2); to obtain enough O_2 by simple diffusion, they must live in an environment with a relatively high oxygen concentration. Bacteria can thrive on 1 percent of the current atmospheric O_2 levels; eukaryotic cells require levels that are at least 2–3 percent of current concentrations. (For concentrations of dissolved O_2 in the oceans to reach these levels, much higher atmospheric concentrations were needed.)

Probably because it took many millions of years for Earth to develop an oxygenated atmosphere, only unicellular prokaryotes lived on Earth for more than 2 billion years. About 1.5 bya, atmospheric O_2 concentrations became high enough for large eukaryotic cells to flourish (Figure 25.5). Further increases in atmospheric O_2 levels 750 to 570 million years ago (mya) enabled several groups of multicellular organisms to evolve.



25.5 Larger Cells, Larger Organisms Need More Oxygen Changes in oxygen concentrations have strongly influenced, and been influenced by, the evolution of life. (Note that the horizontal axis of the graph is on a logarithmic scale.)

INVESTIGATING LIFE

25.6 Rising Oxygen Levels and Body Size in Insects

In this experiment, flies were raised under hyperbaric conditions (increased atmospheric pressure), thus increasing the partial pressure of O₂ in a manner that simulated the greater levels of atmospheric O₂ characteristic of the Carboniferous and Permian. Robert Dudley asked if flies raised in hyperbaric conditions would grow larger than their normal counterparts.

HYPOTHESIS

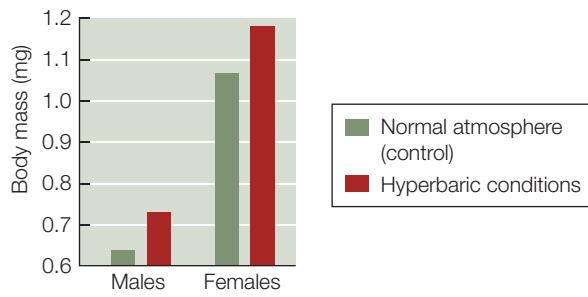
Under increased atmospheric pressure, the increased partial pressure of O₂ will allow directional selection for increased body size in flying insects.

METHOD

- Divide a population of fruit flies (*Drosophila melanogaster*) into two lines.
- Raise one line (the control) at current atmospheric oxygen conditions. Raise the experimental line in hyperbaric conditions (increased partial pressure of O₂, simulating increased atmospheric oxygen concentrations). Continue for 5 generations.
- Raise the F₆ offspring of both lines under identical environmental conditions.
- Weigh all the F₆ individuals and test for statistical differences in the average body mass of the flies in each population.

RESULTS

The average body mass of F₆ individuals of both sexes in the experimental line was significantly ($p < 0.0001$) greater than that of insects in the control line.



CONCLUSION

In at least some flying insects, increased concentrations of oxygen could lead to a long-term evolutionary trend toward increased body size.

FURTHER INVESTIGATION: How would you confirm that the change in average body size is related to increased partial pressure of O₂, and not to other aspects of overall increased atmospheric pressure?

Go to yourBioPortal.com for original citations, discussions, and relevant links for all INVESTIGATING LIFE figures.

O₂ concentrations increased again during the Carboniferous and Permian periods because of the evolution of large vascular plants in the expansive lowland swamps that existed then (see

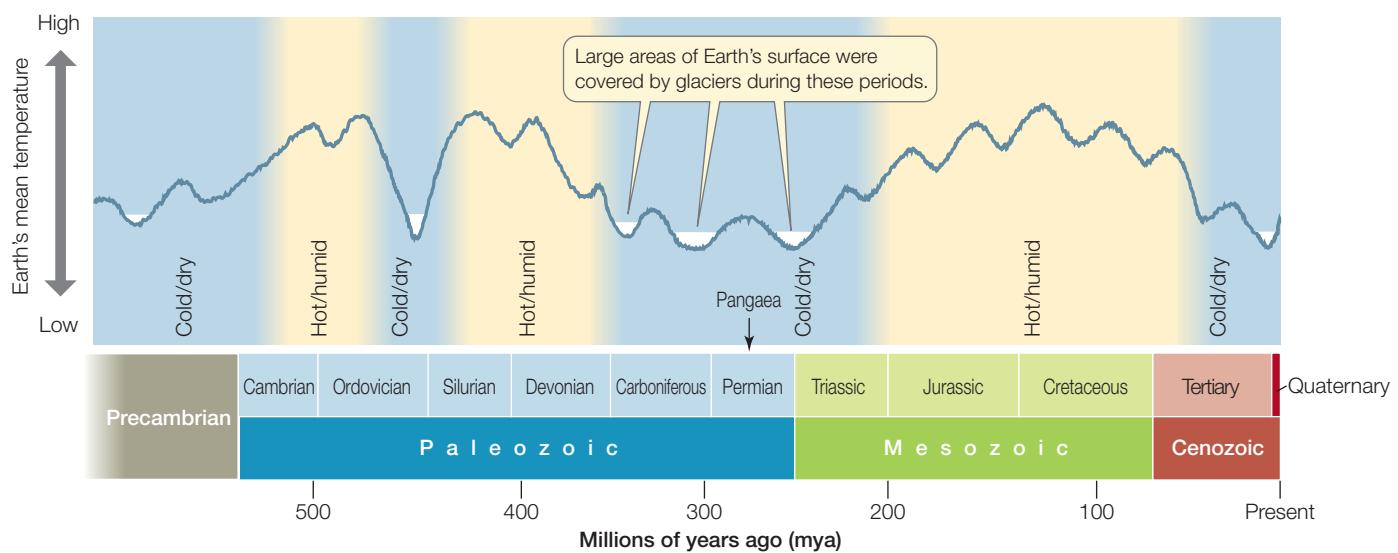
Table 25.1). These swamps resulted in extensive burial of plant debris from vascular plants, which led to the formation of Earth's vast coal deposits. As the buried organic material was not subject to oxidation, and the living plants were producing large quantities of O₂, atmospheric O₂ increased to concentrations that have not been reached again in Earth's history (see Figure 25.5). As mentioned in the opening of this chapter, high concentrations of atmospheric O₂ allowed the evolution of giant flying insects and amphibians that could not survive in today's atmosphere. The drying of the lowland swamps at the end of the Permian reduced global organic burial, and also the production of atmospheric O₂, so O₂ concentrations dropped rapidly. Over the past 200 million years, with the diversification of flowering plants, O₂ concentrations have again increased, but not to the levels that characterized the Carboniferous and Permian periods.

Biologists have conducted experiments that demonstrate the changing selective pressures that can accompany changes in O₂ levels. In experimental conditions, an increase in O₂ concentration can be simulated by increasing atmospheric pressure in a hyperbaric chamber. Increasing atmospheric pressure increases the *partial pressure of oxygen* (see Chapter 49) in a manner that simulates an increase in O₂ concentration at normal atmospheric pressure. When lines of fruit flies (*Drosophila*) are raised in artificial hyperbaric atmospheres (which have higher partial pressure of O₂), they quickly evolve larger body sizes over just a few generations (Figure 25.6). The current levels of atmospheric O₂ appear to constrain body size evolution of these flying insects; increases in O₂ appear to relax these constraints. This demonstrates that the *stabilizing selection* on body size at present O₂ concentrations can quickly switch to *directional selection* (see Section 21.3) for a change in body size in response to a change in O₂ levels. Directional selection over a period of millions of years would be sufficient to account for giant insects such as *Meganeuropsis*, described at the beginning of this chapter.

Many physical conditions on Earth have oscillated in response to the planet's internal processes, such as volcanic activity and continental drift. Extraterrestrial events, such as collisions with meteorites, have also left their mark. In some cases, as we saw earlier and will see again in this chapter, changing physical parameters caused **mass extinctions**, during which a large proportion of the species living at the time disappeared. After each mass extinction, the diversity of life rebounded, but recovery took millions of years.

Earth's climate has shifted between hot/humid and cold/dry conditions

Through much of its history, Earth's climate was considerably warmer than it is today, and temperatures decreased more gradually toward the poles. At other times, Earth was colder than it is today. Large areas were covered with glaciers near the end of the Precambrian and Ordovician, and during parts of the Carboniferous and Permian periods. These cold periods were separated by long periods of milder climates (Figure 25.7). Because we are living in one of the colder periods in Earth's history, it is difficult for us to imagine the mild climates that were found at high latitudes during much of the history of life. During the



25.7 Hot/Humid and Cold/Dry Conditions Have

Alternated over Earth's History Throughout Earth's history, periods of cold climates and glaciations (white depressions) have been separated by long periods of milder climates.

In the Quaternary period there has been a series of glacial advances, interspersed with warmer interglacial intervals during which the glaciers retreated.

"Weather" refers to daily events, such as individual storms. "Climate" refers to long-term average expectations of the various seasons at a given location. Weather often changes rapidly; climates typically change slowly. Major climatic shifts have taken place over periods as short as 5,000 to 10,000 years, primarily as a result of changes in Earth's orbit around the sun. A few climatic shifts have been even more rapid. For example, during one Quaternary interglacial period, the ice-locked Antarctic Ocean became nearly ice-free in less than 100 years. Such rapid changes are usually caused by sudden shifts in ocean currents. Some climate changes have been so rapid that the extinctions caused by them appear to be nearly "instantaneous" in the fossil record.

We are currently living in a time of rapid climate change thought to be caused by a buildup of atmospheric CO₂, primarily from the burning of fossil fuels. We are reversing the process of organic burial that occurred (especially) in the Carboniferous and Permian, but we are doing so over a few hundred years rather than the many millions of years over which these deposits accumulated. The current rate of increase of atmospheric CO₂ is unprecedented in Earth's history. A doubling of the atmospheric CO₂ concentration—which may happen during the current century—is expected to increase the average temperature of Earth, change rainfall patterns, melt glaciers and ice caps, and raise sea level. The possible consequences of such climate changes are discussed in Chapters 58 and 59.

Volcanoes have occasionally changed the history of life

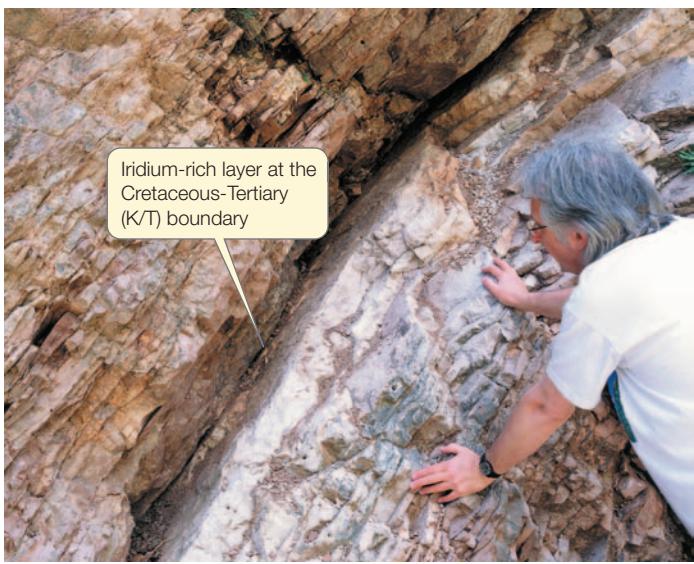
Most volcanic eruptions produce only local or short-lived effects, but a few large volcanic eruptions have had major consequences for life. When Krakatoa erupted in Indonesia in 1883,

it ejected more than 25 cubic kilometers of ash and rock, as well as large quantities of sulphur dioxide gas (SO₂). The SO₂ was ejected into the stratosphere and then moved by high-level winds around the planet. This led to high concentrations of sulphurous acid (H₂SO₃) in high-level clouds, which meant less sunlight got through to Earth's surface. Global temperatures dropped by 1.2°C in the year following the eruption, and global weather patterns showed strong effects for another 5 years. This was all the result of a single volcanic eruption. The collision of continents during the Permian period (about 275 mya) formed a single, gigantic land mass (Pangaea) and caused many massive volcanic eruptions. These eruptions resulted in considerable blockage of sunlight, contributing to the glaciations of that time (see Figure 25.7). Massive volcanic eruptions occurred again as the continents drifted apart during the late Triassic and at the end of the Cretaceous.

Extraterrestrial events have triggered changes on Earth

At least 30 meteorites between the sizes of baseballs and soccer balls hit Earth each year. Collisions with large meteorites or comets are rare, but such collisions have probably been responsible for several mass extinctions. Several types of evidence tell us about these collisions. Their craters, and the dramatically disfigured rocks that resulted from their impact, are found in many places. Geologists have also discovered compounds in these rocks that contain helium and argon with isotope ratios characteristic of meteorites, which are very different from the ratios found elsewhere on Earth.

A meteorite caused or contributed to a mass extinction at the end of the Cretaceous period (about 65 mya). The first clue that a meteorite was responsible came from the abnormally high concentrations of the element iridium in a thin layer separating rocks deposited during the Cretaceous from those deposited during the Tertiary (Figure 25.8). Iridium is abundant in some meteorites, but it is exceedingly rare on Earth's surface. Scientists discovered a circular crater 180 kilometers in diameter buried beneath the northern coast of the Yucatán Peninsula of



25.8 Evidence of a Meteorite Impact The white layers of rock are Cretaceous in age; the layers at the upper left were deposited in the Tertiary. Between the two is a thin, dark layer of clay that contains large amounts of iridium, a metal common in some meteorites but rare on Earth. Its high concentration in sediments deposited about 65 million years ago suggests the impact of a large meteorite.

Mexico. When it collided with Earth, the meteorite released energy equivalent to that of 100 million megatons of high explosives, creating great tsunamis. A massive plume of debris swelled to a diameter of up to 200 kilometers, spread around Earth, and descended. The descending debris heated the atmosphere to several hundred degrees, ignited massive fires, and blocked the sun, preventing plants from photosynthesizing. The settling debris formed the iridium-rich layer. About a billion tons of soot, which has a composition that matches smoke from forest fires, was also deposited. Many fossil species (particularly dinosaurs) that are found in Cretaceous rocks are not found in the Tertiary rocks of the next layer.

25.2 RECAP

Conditions on Earth have changed dramatically over time. Changes in atmospheric concentrations of O₂ and in Earth's climate have had major effects on biological evolution. Continental drift, volcanic eruptions, and large meteorite strikes have contributed to climatic changes during Earth's history.

- Describe how increases in atmospheric concentrations of O₂ affected the evolution of multicellular organisms. **See pp. 523–524 and Figure 25.5**
- How have volcanic eruptions and meteorite strikes influenced the course of life's evolution? **See p. 525**

The many dramatic physical events of Earth's history have influenced the nature and timing of evolutionary changes among Earth's living organisms. We now will look more closely at some of the major events that characterize the history of life on Earth.

25.3 What Are the Major Events in Life's History?

Life first evolved on Earth about 3.8 bya. By about 1.5 bya, eukaryotic organisms had evolved (see Table 25.1). The fossil record of organisms that lived prior to 550 mya is fragmentary, but it is good enough to show that the total number of species and individuals increased dramatically in late Precambrian times. As discussed above, pre-Darwinian geologists divided geological history into eras and periods based on their distinct fossil assemblages. Biologists refer to the assemblage of all organisms of all kinds living at a particular time or place as a **biota**. All of the plants living at a particular time or place are its **flora**; all of the animals are its **fauna**. Table 25.1 describes some of the physical and biological changes, such as mass extinctions and dramatic increases in the diversity of major groups of organisms, associated with each unit of time.

About 300,000 species of fossil organisms have been described, and the number steadily grows. The number of named species, however, is only a tiny fraction of the species that have ever lived. We do not know how many species lived in the past, but we have ways of making reasonable estimates. Of the present-day biota, nearly 1.8 million species have been named. The actual number of living species is probably well over 10 million, and possibly much higher, because many species have not yet been discovered and described by biologists. So the number of described fossil species is only about 3 percent of the estimated minimum number of living species. Life has existed on Earth for about 3.8 billion years. Many species last only a few million years before undergoing speciation or going extinct; therefore, Earth's biota must have turned over many times during geological history. So the total number of species that have lived over evolutionary time must vastly exceed the number living today. Why have only about 300,000 of these tens of millions of species been described from fossils to date?

Several processes contribute to the paucity of fossils

Only a tiny fraction of organisms ever become fossils, and only a tiny fraction of fossils are ever discovered by paleontologists. Most organisms live and die in oxygen-rich environments in which they quickly decompose. They are not likely to become fossils unless they are transported by wind or water to sites that lack oxygen, where decomposition proceeds slowly or not at all. Furthermore, geological processes often transform rocks, destroying the fossils they contain, and many fossil-bearing rocks are deeply buried and inaccessible. Paleontologists have studied only a tiny fraction of the sites that contain fossils, but they find and describe many new ones every year.

The fossil record is most complete for marine animals that had hard skeletons (which resist decomposition). Among the nine major animal groups with hard-shelled members, approximately 200,000 species have been described from fossils—roughly twice the number of living marine species in these same groups. Paleontologists lean heavily on these groups in their interpretations of the evolution of life. Insects and spiders are also relatively well represented in the fossil record, because they are numerically abundant and have hard exoskeletons



Solenopsis sp.

25.9 Insect Fossils Chunks of amber—fossilized tree resin—often contain insects that were preserved when they were trapped in the sticky resin. This fire ant fossil is some 30 million years old.

(Figure 25.9). The fossil record, though incomplete, is good enough to document clearly the factual history of the evolution of life.

By combining information about geological changes during Earth's history with evidence from the fossil record, scientists have composed portraits of what Earth and its inhabitants may have looked like at different times. We know in general where the continents were and how life changed over time, but many of the details are poorly known, especially for events in the more remote past.

Precambrian life was small and aquatic

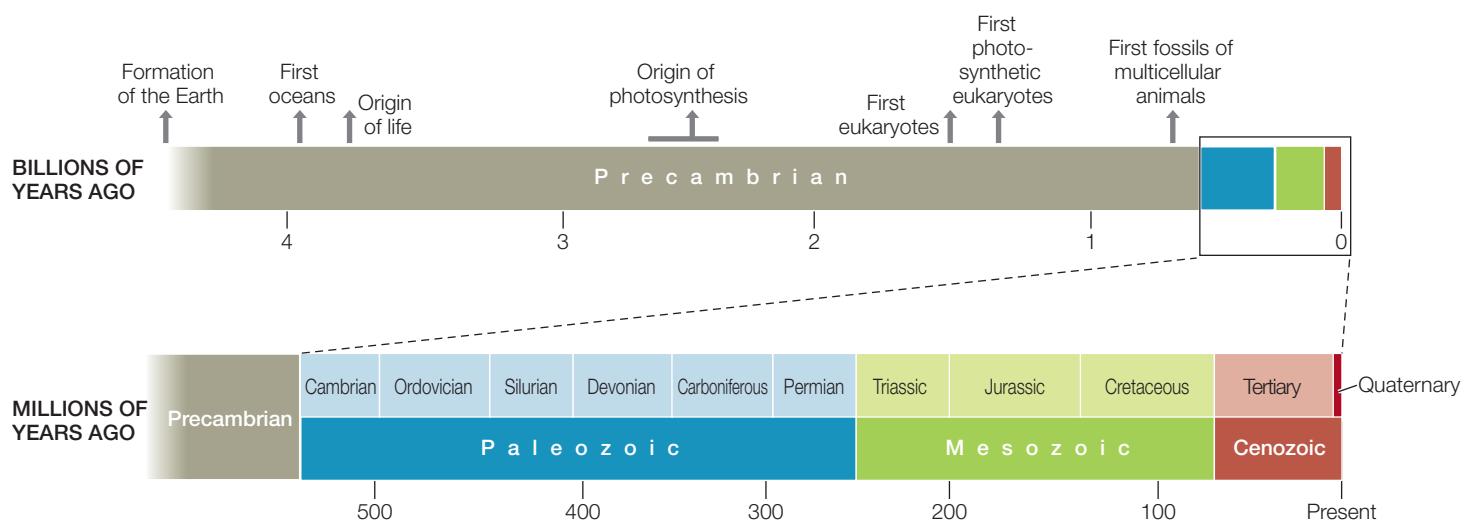
For most of its history, life was confined to the oceans, and all organisms were small. Over the long ages of the **Precambrian era**—more than 3 billion years—the shallow seas slowly be-

gan to teem with life. For most of the Precambrian, life consisted of microscopic prokaryotes; eukaryotes evolved about two-thirds of the way through the era (**Figure 25.10**). Unicellular eukaryotes and small multicellular animals fed on floating photosynthetic microorganisms. Small floating organisms, known collectively as *plankton*, were strained from the water and eaten by slightly larger *filter-feeding* animals. Other animals ingested sediments on the seafloor and digested the remains of organisms within them. By the late Precambrian (630–542 mya), many kinds of multicellular soft-bodied animals had evolved. Some of them were very different from any animals living today, and may be members of groups that have no living descendants (**Figure 25.11**).

Life expanded rapidly during the Cambrian period

The Cambrian period (542–488 mya) marks the beginning of the Paleozoic era. The oxygen concentration in the Cambrian atmosphere was approaching its current level, and the land masses had come together to form several large continents. A geologically rapid diversification of life took place that is sometimes referred to as the **Cambrian explosion** (although in fact it began before the Cambrian, and the “explosion” took millions of years). Several of the major groups of animals that have species living today first evolved during the Cambrian. An overview of the continental and biotic shifts that characterized the Cambrian and subsequent periods is shown in **Figure 25.12** on the following pages.

For the most part, fossils tell us only about the hard parts of organisms, but in three known Cambrian fossil beds—the Burgess Shale in British Columbia, Sirius Passet in northern Greenland, and the Chengjiang site in southern China—the soft parts of many animals were preserved. Crustacean arthropods (crabs, shrimps, and their relatives) are the most diverse group



25.10 A Sense of Life's Time The top timeline shows the 4.5 billion year history of life on Earth. Most of this time is accounted for by the Precambrian, a 3.4 billion year era that saw the origin of life and the evolution of cells, photosynthesis, and multicellularity. The final 600 million years are expanded in the second timeline and detailed in Figure 25.12.



Spriggina floundersi



Mawsonites spriggi



Dickinsonia costata

25.11 Precambrian Life These fossils of soft-bodied invertebrates, excavated at Ediacara in southern Australia, were formed about 600 million years ago. Very different from later life forms, they illustrate the diversity of life at the end of the Precambrian era.

in the Chinese fauna; some of them were large carnivores. Multicellular diversity was largely or completely aquatic during the Cambrian. If there was life on land at this time, it was probably restricted to microbial organisms.

Many groups of organisms that arose during the Cambrian later diversified

Geologists divide the remainder of the Paleozoic era into the Ordovician, Silurian, Devonian, Carboniferous, and Permian periods. Each period is characterized by the diversification of specific groups of organisms. Mass extinctions marked the ends of the Ordovician, Devonian, and Permian.

THE ORDOVICIAN (488–444 MYA) During the Ordovician period, the continents, which were located primarily in the Southern Hemisphere, still lacked multicellular plants. Evolutionary radiation of marine organisms was spectacular during the early Ordovician, especially among animals, such as brachiopods and mollusks, that lived on the seafloor and filtered small prey from the water. At the end of the Ordovician, as massive glaciers formed over the southern continents, sea levels dropped about 50 meters and ocean temperatures dropped. About 75 percent of the animal species became extinct, probably because of these major environmental changes.

THE SILURIAN (444–416 MYA) During the Silurian period, the continents began to merge together. Marine life rebounded from the mass extinction at the end of the Ordovician. Animals able to swim in open water and feed above the ocean bottom appeared for the first time. Jawless fishes diversified, and the first ray-finned fishes evolved. The tropical sea was uninterrupted by land barriers, and most marine organisms were widely distributed. On land, the first vascular plants evolved late in the Silurian (about 420 mya). The first terrestrial arthropods—scorpions and millipedes—evolved at about the same time.

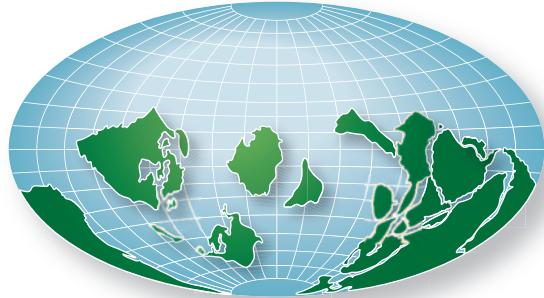
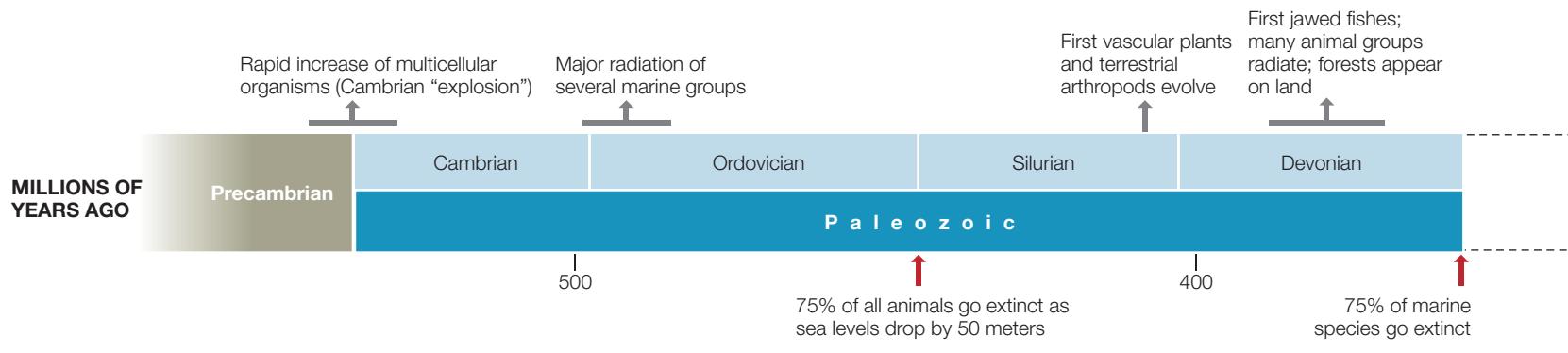
THE DEVONIAN (416–359 MYA) Rates of evolutionary change accelerated in many groups of organisms during the Devonian period. The major land masses continued to move slowly toward each other. In the oceans there were great evolutionary radiations of corals and of shelled, squidlike cephalopod mollusks. Fishes diversified as jawed forms replaced jawless ones and as heavy armor gave way to the less rigid outer coverings of modern fishes.

Terrestrial communities changed dramatically during the Devonian. Club mosses, horsetails, and tree ferns became common; some attained the size of large trees. Their roots accelerated the weathering of rocks, resulting in the development of the first forest soils. The ancestors of gymnosperms—the first plants to produce seeds—appeared in the Devonian. The first known fossils of centipedes, spiders, mites, and insects date to this period, and fishlike amphibians began to occupy the land.

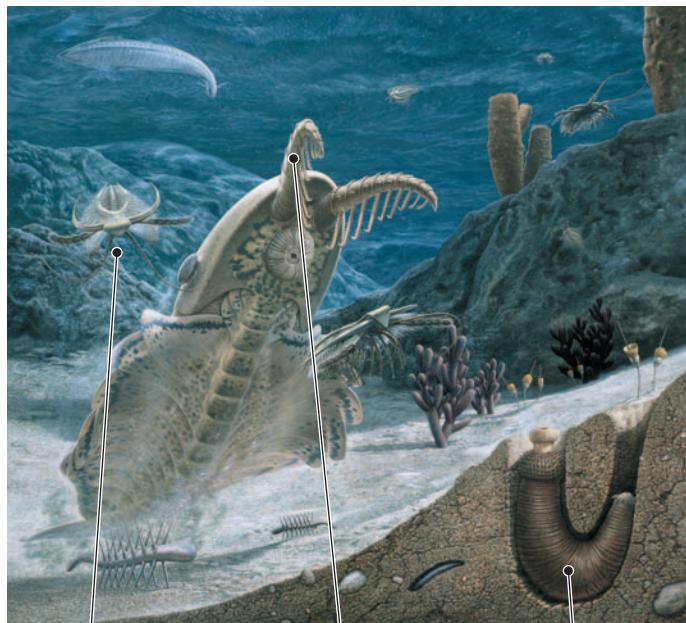
A massive extinction of about 75 percent of all marine species marked the end of the Devonian. Paleontologists are uncertain about its cause, but two large meteorites that collided with Earth at that time (one in present-day Nevada and the other in western Australia) may have been responsible, or at least a contributing factor. The continued coalescence of the continents, with the corresponding reduction in continental shelves, may have also contributed to this mass extinction event.

THE CARBONIFEROUS (359–297 MYA) Large glaciers formed over high-latitude portions of the southern land masses during the Carboniferous period, but extensive swamp forests grew on the tropical continents. These forests were not made up of the kinds of trees we know today, but were dominated by giant tree ferns and horsetails with small leaves. Fossilized remains of those forests formed the coal we now mine for energy. In the seas,

25.12 A Brief History of Life on Earth The geologically rapid “explosion” of life during the Cambrian saw the rise of several animal groups that have representatives surviving today. The following three pages depict life’s history from the Cambrian forward. Movements of the major continents during the past half-billion years are shown in the maps of Earth, and associated biotas for each time period are depicted. The artists’ reconstructions are based on fossils such as those shown in the photographs.



Cambrian



Marrella splendens



Ottoia sp.



Anomalocaris canadensis (claw only)



Devonian



Codiocrinus schultzei



Phacops ferdinandi



Eridophyllum sp.



Orthoconic nautiloid

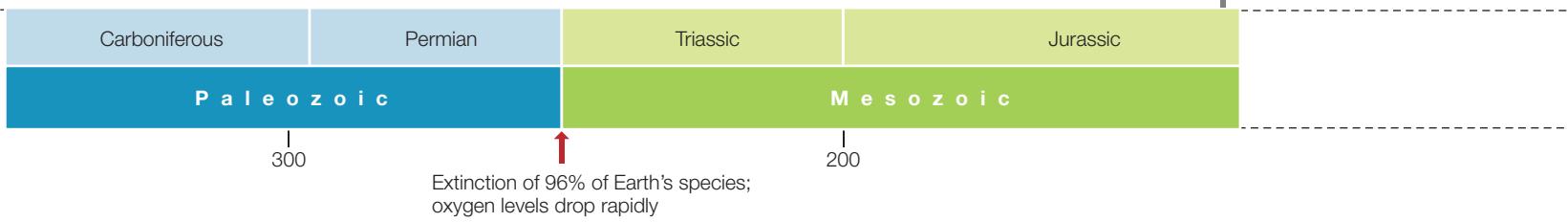
Extensive swamp forests produce coal; origin of amniotes; great increase in terrestrial animal diversity

Giant amphibians and flying insects; ray-finned fishes abundant in freshwater

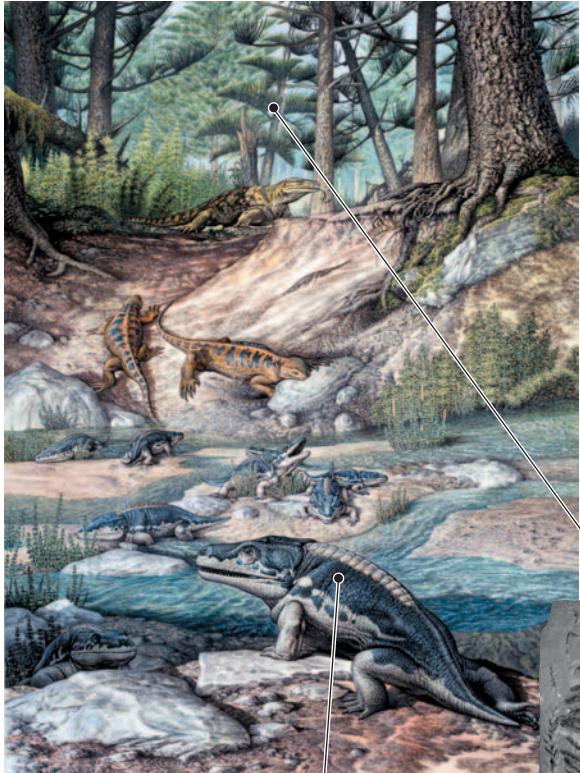
On land, conifers become dominant plants; frogs and reptiles begin to diversify

Dinosaurs, pterosaurs, ray-finned fishes diversify; first mammals appear

First known flowering plant fossils



Permian



Walchia piniformis

Cacops sp.

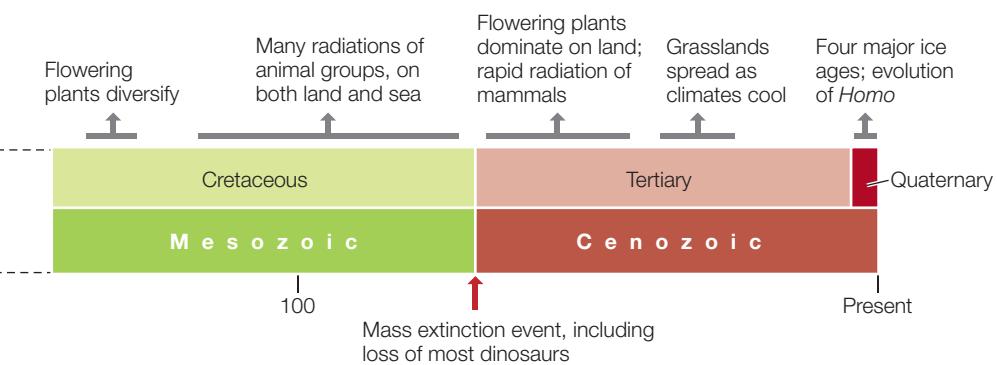
Triassic



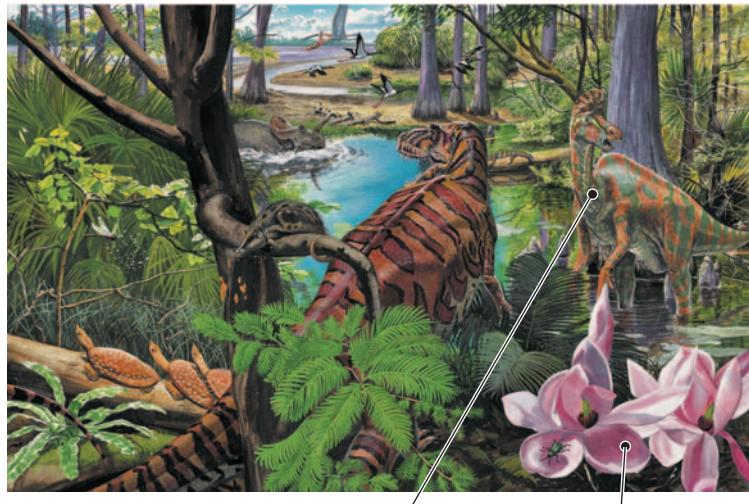
Phlebopteris smithii

Coelophysis bauri

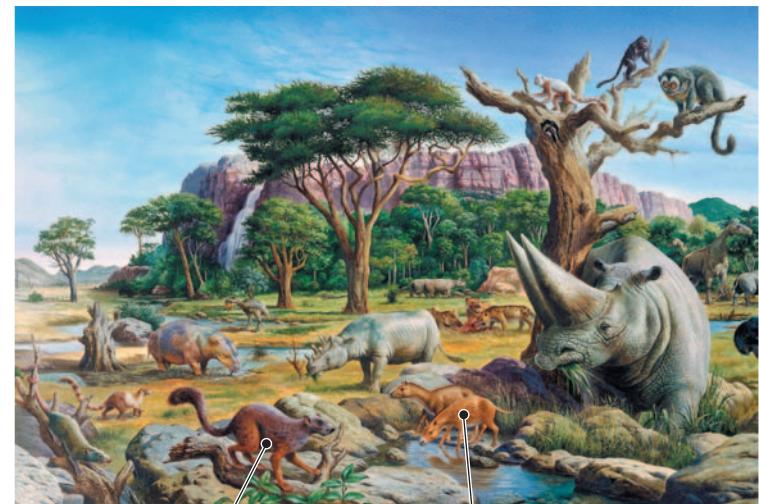




Cretaceous



Tertiary



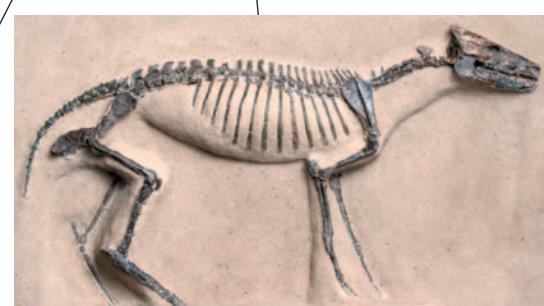
Gryposaurus sp.



Magnolia sp.



Plesiadapis fodinatus (jaw)



Hyracotherium leporinum

25.13 Evidence of Insect Diversification

The margins of this fossil fern leaf from the Carboniferous have been chewed by insects.



crinoids (sea lilies and feather stars) reached their greatest diversity, forming “meadows” on the seafloor.

The diversity of terrestrial animals increased greatly during the Carboniferous. Snails, scorpions, centipedes, and insects were abundant and diverse. Insects evolved wings, becoming the first animals to fly. Flight gave herbivorous insects easy access to tall plants; plant fossils from this period show evidence of chewing by insects (Figure 25.13). The terrestrial vertebrate lineage split, and amphibians became larger and better adapted to terrestrial existence, while the sister lineage led to the **amniotes**, vertebrates with well-protected eggs that can be laid in dry places.

THE PERMIAN (297–251 MYA) During the Permian period, the continents coalesced completely into the supercontinent **Pangaea**. Permian rocks contain representatives of many major groups of insects we know today. By the end of the period, the reptiles split from a second amniote lineage (which would lead to the mammals). Ray-finned fishes became common in the fresh waters of Pangaea.

Conditions for life deteriorated toward the end of the Permian. Massive volcanic eruptions resulted in outpourings of lava that covered large areas of Earth. The ash and gases produced by the volcanoes blocked sunlight and cooled the climate, resulting in the largest glaciers in Earth’s history. Atmospheric oxygen concentrations gradually dropped from about 30 to 15 percent. At such low concentrations, most animals would have been unable to survive at elevations above 500 meters; thus about half of the land area would have been uninhabitable at the end of the Permian. The combination of these changes resulted in the most drastic mass extinction event in Earth’s history. Scientists estimate that about 96 percent of all species became extinct at the end of the Permian.

Geographic differentiation increased during the Mesozoic era

The few organisms that survived the Permian mass extinction found themselves in a relatively empty world at the start of the

Mesozoic era (251 mya). As Pangaea slowly began to break apart, the oceans rose and once again flooded the continental shelves, forming huge, shallow inland seas. Atmospheric oxygen concentrations gradually rose. Life once again proliferated and diversified, but different groups of organisms came to the fore. The three groups of phytoplankton (floating photosynthetic organisms) that dominate today’s oceans—dinoflagellates, coccolithophores, and diatoms—became ecologically important at this time; their remains are the primary origin of the world’s oil deposits. Seed-bearing plants replaced the trees that had ruled the Permian forests.

The Mesozoic era is divided into three periods: the Triassic, Jurassic, and Cretaceous. The Triassic and Cretaceous were terminated by mass extinctions, probably caused by meteorite impacts.

THE TRIASSIC (251–200 MYA) Pangaea began to break apart during the Triassic period. Many invertebrate groups became more species-rich, and many burrowing animals evolved from groups living on the surfaces of seafloor sediments. On land, conifers and seed ferns were the dominant trees. The first frogs and turtles appeared. A great radiation of reptiles began, which eventually gave rise to crocodilians, dinosaurs, and birds. The end of the Triassic was marked by a mass extinction that eliminated about 65 percent of the species on Earth.

THE JURASSIC (200–145 MYA) During the Jurassic period, Pangaea became fully divided into two large continents: **Laurasia** drifted northward and **Gondwana** drifted south. Ray-finned fishes rapidly diversified in the oceans. The first lizards appeared, and flying reptiles (pterosaurs) evolved. Most of the large terrestrial predators and herbivores of the period were dinosaurs. Several groups of mammals made their first appearance, and the earliest known fossils of flowering plants are from late in this period.

THE CRETACEOUS (145–65 MYA) By the early Cretaceous period, Laurasia and Gondwana had begun to break apart into the con-

tinents we know today. A continuous sea encircled the tropics. Sea levels were high, and Earth was warm and humid. Life proliferated both on land and in the oceans. Marine invertebrates increased in diversity and in number of species. On land, the reptile radiation continued as dinosaurs diversified further and the first snakes appeared. Early in the Cretaceous, flowering plants began the radiation that led to their current dominance of the land. By the end of the period, many groups of mammals had evolved. Most early mammals were small, but one species recently discovered in China, *Repenomamus gigantus*, was large enough to capture and eat young dinosaurs.

As described in Section 25.2, another meteorite-caused mass extinction took place at the end of the Cretaceous (the impact site was near the present day Yucatán Peninsula of Mexico). In the seas, many planktonic organisms and bottom-dwelling invertebrates became extinct. On land, almost all animals larger than about 25 kilograms in body weight became extinct. Many species of insects died out, perhaps because the growth of their food plants was greatly reduced following the impact. Some species in northern North America and Eurasia survived in areas that were not subjected to the devastating fires that engulfed most low-latitude regions.

Modern biota evolved during the Cenozoic era

By the early Cenozoic era (65 mya), the positions of the continents resembled those of today, but Australia was still attached to Antarctica, and the Atlantic Ocean was much narrower. The Cenozoic was characterized by an extensive radiation of mammals, but other groups were also undergoing important changes.

Flowering plants diversified extensively and came to dominate world forests, except in the coolest regions, where the forests were composed primarily of gymnosperms. Mutations of two genes in one group of plants (the legumes) allowed them to use atmospheric nitrogen directly by forming symbioses with a few species of nitrogen-fixing bacteria (see Section 36.4). The evolution of this symbiosis between certain early Cenozoic plants and these specialized bacteria was the first “green revolution” and dramatically increased the amount of nitrogen available for terrestrial plant growth; the symbiosis remains fundamental to the ecological base of life as we know it today.

The Cenozoic era is divided into the Tertiary and the Quaternary periods. Because both the fossil record and our subsequent knowledge of evolutionary history become more extensive as we approach our own time, paleontologists have subdivided these two periods into *epochs* (Table 25.2).

THE TERTIARY (65–2.6 MYA) During the Tertiary period, Australia began its northward drift. By 20 mya it had nearly reached its current position. The early Tertiary was a hot and humid time, and the ranges of many plants shifted latitudinally. The tropics were probably too hot for rainforests and were clothed in low-lying vegetation instead. In the middle of the Tertiary, however, Earth's climate became considerably cooler and drier. Many lineages of flowering plants evolved herbaceous (nonwoody) forms, and grasslands spread over much of Earth.

TABLE 25.2

Subdivisions of the Cenozoic Era

PERIOD	EPOCH	ONSET (MYA)
Quaternary	Holocene ^a	0.01 (~10,000 years ago)
	Pleistocene	2.6
Tertiary	Pliocene	5.3
	Miocene	23
	Oligocene	34
	Eocene	55.8
	Paleocene	65

^aThe Holocene is also known as the Recent.

By the start of the Cenozoic era, invertebrate faunas had already come to resemble those of today. It is among the terrestrial vertebrates that evolutionary changes during the Tertiary were most rapid. Frogs, snakes, lizards, birds, and mammals all underwent extensive radiations during this period. Three waves of mammals dispersed from Asia to North America across one of the several land bridges that have intermittently connected the two continents during the past 55 million years. Rodents, marsupials, primates, and hoofed mammals appeared in North America for the first time.

THE QUATERNARY (2.6 MYA TO PRESENT) We are living in the Quaternary period. It is subdivided into two epochs, the Pleistocene and the Holocene (the Holocene also being known as the Recent).

The Pleistocene was a time of drastic cooling and climate fluctuations. During 4 major and about 20 minor “ice ages,” massive glaciers spread across the continents, and the ranges of animal and plant populations shifted toward the equator. The last of these glaciers retreated from temperate latitudes less than 15,000 years ago. Organisms are still adjusting to these changes. Many high-latitude ecological communities have occupied their current locations for no more than a few thousand years.

It was during the Pleistocene that divergence within one group of mammals, the primates, resulted in the evolution of the hominoid lineage. Subsequent hominoid radiation eventually led to the species *Homo sapiens*—modern humans (see Section 33.5). Many large bird and mammal species became extinct in Australia and in the Americas when *H. sapiens* arrived on those continents about 45,000 and 15,000 years ago, respectively. Many paleontologists believe these extinctions were probably the result of hunting and other influences of *Homo sapiens*.

The tree of life is used to reconstruct evolutionary events

The fossil record reveals broad patterns in life's evolution. To reconstruct major events in the history of life, biologists also rely on the phylogenetic information in the tree of life (see Chapter 22 and the Tree of Life Appendix). We can use phylogeny (in combination with the paleontological record) to reconstruct the

timing of such major events as the acquisition of mitochondria in the ancestral eukaryotic cell, the several independent origins of multicellularity, and the movement of life onto dry land. We can also follow major changes in the genomes of organisms, and even reconstruct many gene sequences of species that are long extinct (see Chapter 24).

Changes to the physical environment on Earth have clearly influenced the great diversity in living organisms we see on the planet today. To study the evolution of that diversity, biologists examine the evolutionary relationships among species. Deciphering these relationships is an important step in understanding how life has diversified on Earth. Part Seven of this book explores the major groups of life and the different solutions that have evolved for major functions such as reproduction, energy acquisition, dispersal, and escape from predation.

[yourBioPortal.com](#)
GO TO The Interactive Tree of Life

25.3 RECAP

Life evolved in the Precambrian oceans. It diversified as atmospheric oxygen approached its current level and the continents came together to form several large land masses. Numerous climate changes and rearrangements of the continents, as well as meteorite impacts, contributed to five major mass extinctions.

- Why have so few of the multitudes of organisms that have existed over millennia become fossilized? [See pp. 526–527](#)
- What do we mean when we refer to the “Cambrian explosion”? [See p. 527](#)
- In what ways has continental drift affected the evolution of life on Earth? [See Figure 25.12](#)

CHAPTER SUMMARY

25.1 How Do Scientists Date Ancient Events?

- The relative ages of organisms can be determined by the dating of fossils and the **strata** of **sedimentary rocks** in which they are found.
- Paleontologists use a variety of radioisotopes with different **half-lives** to date events at different times in the remote past. [Review Figure 25.1](#)
- Geologists divide the history of life into eras and periods, based on major differences in the fossil assemblages found in successive layers of rocks. [Review Table 25.1](#)

25.2 How Have Earth’s Continents and Climates Changed over Time?

- Earth’s crust consists of solid lithospheric plates that float on fluid magma. **Continental drift** caused by convection currents in the magma moves these plates and the continents that lie on top of them. [Review Figure 25.2, ANIMATED TUTORIAL 25.1](#)
- Conditions on Earth have changed dramatically over time. Increases in atmospheric oxygen and changes in Earth’s climate have greatly influenced the evolution of life on Earth. [Review Figures 25.5 and 25.7](#)
- Oxygen-generating cyanobacteria liberated enough O₂ to open the door to oxidation reactions in metabolic pathways. The aerobic prokaryotes were able to harvest more energy than anaerobic organisms and began to proliferate. Increases in atmospheric O₂ levels supported the evolution of large eukaryotic cells.

- Major physical events on Earth, such as the collision of continents that formed the supercontinent **Pangaea**, have affected Earth’s surface, climate, and atmosphere. In addition, extraterrestrial events such as meteorite strikes created sudden and dramatic environmental shifts. All of these changes have affected the history of life.

25.3 What Are The Major Events in Life’s History?

- Paleontologists use fossils and evidence of geological changes to determine what Earth and its **biota** may have looked like at different times.
- During most of its history, life was confined to the oceans. Multicellular life diversified extensively during the **Cambrian explosion**. [Review Figure 25.11](#)
- The periods of the Paleozoic era were each characterized by the diversification of specific groups of organisms. **Amniotes**—vertebrates whose eggs can be laid in dry places—first appeared during the Carboniferous period.
- During the Mesozoic era, distinct terrestrial biotas evolved on each continent.
- Five episodes of **mass extinction** punctuated the history of life in the Paleozoic and Mesozoic eras.
- Earth’s **flora** has been dominated by flowering plants since the Cenozoic era.
- Phylogenetic trees help reconstruct the timing of evolutionary events and clarify relationships among modern species.

SEE WEB ACTIVITY 25.1 for a concept review of this chapter.

SELF-QUIZ

1. Which of the following is *not* true of the giant flying dragonflies of the Carboniferous and Permian?
 - a. Some species grew to have wing spans as wide or wider than many modern birds of prey.
 - b. They were the largest flying predators of the time.
 - c. Such large flying insects could exist because of the higher concentrations of atmospheric oxygen compared to the present.
 - d. Their predators were giant reptiles.
 - e. Fossils of one large species, *Meganeurosis permiana*, have been found in the Permian rocks of Kansas.
2. In undisturbed strata of sedimentary rock, the oldest rocks
 - a. lie at the top.
 - b. lie at the bottom.
 - c. are in the middle.
 - d. are distributed among the strata of younger rocks.
 - e. none of the above
3. ^{14}C can be used to determine the ages of fossil organisms because
 - a. all organisms contain many carbon compounds.
 - b. ^{14}C has a regular rate of decay to ^{14}N .
 - c. the ratio of ^{14}C to ^{12}C in living organisms is always the same as that in the atmosphere.
 - d. the production of new ^{14}C in the atmosphere just balances the natural radioactive decay of ^{14}C .
 - e. all of the above
4. The concentration of oxygen in the Earth's atmosphere
 - a. has increased steadily through time.
 - b. has decreased steadily through time.
 - c. has been both higher and lower in the past than at present.
 - d. was lower during most of the Permian than at present.
 - e. was at its highest levels in the Cambrian.
5. The total of all species of organisms in a given region is known as the region's
 - a. biota.
 - b. flora.
 - c. fauna.
 - d. flora and fauna.
 - e. biogeography.
6. The coal beds we now mine for energy are largely the remains of
 - a. plants that grew in swamps during the Carboniferous period.
 - b. algae that grew in marshes during the Devonian period.
 - c. giant insects and amphibians of the Permian period.
 - d. plants that grew in the oceans during the Carboniferous period.
 - e. none of the above
7. The mass extinction at the end of the Ordovician period was probably caused by
 - a. the collision of Earth with a large meteorite.
 - b. massive volcanic eruptions.
 - c. massive glaciation on the southern continents and associated climatic changes.
 - d. the uniting of all continents to form Pangaea.
 - e. changes in Earth's orbit.
8. The cause of the mass extinction at the end of the Mesozoic era probably was
 - a. continental drift.
 - b. the collision of Earth with a large meteorite.
 - c. changes in Earth's orbit.
 - d. massive glaciation.
 - e. changes in the salt concentration of the oceans.
9. Which of the following times was marked by the largest mass extinction of life in the history of Earth?
 - a. The end of the Cretaceous
 - b. The end of the Devonian
 - c. The end of the Permian
 - d. The end of the Triassic
 - e. The end of the Silurian
10. Paleontologists have subdivided the Cenozoic era into epochs because
 - a. *Homo sapiens* evolved at the start of the Cenozoic.
 - b. the continents had achieved their present positions.
 - c. the number of species stopped increasing at this time.
 - d. our knowledge of the evolutionary events of the Cenozoic is more extensive than for other eras.
 - e. starting with the Cenozoic, the fossil record is no longer a necessary source of information about evolutionary relationships.

FOR DISCUSSION

1. Some groups of organisms have evolved to contain large numbers of species; other groups have produced only a few species. Is it meaningful to consider the former groups more successful than the latter? What does the word "success" mean in evolution?
2. Scientists date ancient events using a variety of methods, but nobody was present to witness or record those events. Accepting those dates requires us to understand the accuracy and appropriateness of indirect measurement techniques. What other basic scientific concepts are also based on the results of indirect measurement techniques?
3. Why is it useful to be able to date past events absolutely as well as relatively?
4. If we are living during one of the cooler periods in Earth's history, why should we be concerned about human activities that are thought to contribute to global climate warming?
5. What conditions may have favored the evolution of multicellular groups of organisms near the end of the Precambrian?
6. In what ways do endosymbiotic events (such as the origin of mitochondria and chloroplasts) complicate the classification of the major groups of life?

ADDITIONAL INVESTIGATION

The experiment in Figure 25.6 showed that body size of insects may evolve quickly following changes in atmospheric oxygen

concentrations. What other experiments could you devise to test the effects of changing atmospheric oxygen?

26

Bacteria and Archaea: The Prokaryotic Domains

Life on a strange planet

It must have been quite a shock when Thomas “Grif” Taylor’s Antarctic exploration team first spotted Blood Falls in 1911. The blood-red falls were certainly a surprise in the snowy, icy terrain. What could possibly cause a red waterfall in Antarctica?

A few million years ago, the Taylor Glacier (which now bears the explorer’s name) moved above a pool of salty water, trapping the pool under 400 meters of ice. The harsh environment in the resulting enclosed subglacial sea could hardly seem more hostile to life. It is extremely cold; there is no light and virtually no oxygen; and salt concentrations are several times higher than seawater. In short, it is not a place one might expect to find a diverse living ecosystem.

Some water is able to seep out of this subglacial sea. This water is stained a dark, rusty red, and it spills from the head of Taylor Glacier to form Blood Falls. Taylor specu-

lated that red algae might account for the red coloration, but in the 1960s geologist Robert Black discovered that the water’s color arises from iron oxides that come from the underlying bedrock. With the methods then available, biologists could not detect any living organisms in the cold, saline, iron-rich water.

A half-century later, biologists were better equipped to study microscopic life in strange places. By then it was also possible to amplify and sequence genes from single microbes, and to place these gene sequences into the framework of the tree of life to identify and classify the microbes. Microbiologist Jill Mikucki and her colleagues used these techniques on water samples from Blood Falls, and reported in 2009 that the falls contain an unusual ecosystem of at least 17 different species of bacteria. The bacteria survive by metabolizing minute amounts of organic matter trapped in the subglacial sea, using sulfate and iron ions as catalysts and electron acceptors.

The presence of living organisms in Blood Falls confirms that it is hard to find a place on or even near the surface of the Earth that does not contain populations of prokaryotes. There are prokaryotes in volcanic vents, in the clouds, in environments as acidic as battery acid or as alkaline as household ammonia. There are species that can survive below the freezing point and above the boiling point



A Splash of Color in a Frozen World of White Antarctica’s Blood Falls is the outflow of a subglacial sea that contains an unusual ecosystem of bacteria that rely on sulfate and iron ions for metabolism.



Prokaryotes Can Take the Heat Entire ecosystems of prokaryotes create the beauty of Morning Glory Pool, a hot spring in Yellowstone National Park. Cyanobacteria impart the “morning glory blue” color. Archaea live in the intensely hot regions of the pool’s interior.

of water. There are more prokaryotes living on and inside our bodies than we have human cells. The prokaryotes are masters of metabolic ingenuity, having developed more ways to obtain energy from the environment than the eukaryotes have. They have been around much longer than other organisms, too.

Prokaryotes are by far the most numerous organisms on Earth. Late in the twentieth century, it became apparent to microbiologists that all prokaryotes are not most closely related to one another. Two prokaryotic lineages diverged early in life’s evolution: Bacteria and Archaea. An early merging between members of these two groups is thought to have given rise to the eukaryotic lineage, which includes humans.

IN THIS CHAPTER we will discuss the distribution of prokaryotes and examine their remarkable metabolic diversity. We will describe the difficulties involved in determining evolutionary relationships among the prokaryotes and will survey the surprising diversity of organisms in each domain. We will discuss how prokaryotes can have enormous influence on their environments. Finally, we will discuss the evolutionary origin and diversity of viruses and their relationship to the rest of life.

CHAPTER OUTLINE

- 26.1** How Did the Living World Begin to Diversify?
- 26.2** What Are Some Keys to the Success of Prokaryotes?
- 26.3** How Can We Resolve Prokaryote Phylogeny?
- 26.4** What Are the Major Known Groups of Prokaryotes?
- 26.5** How Do Prokaryotes Affect Their Environments?
- 26.6** Where Do Viruses Fit into the Tree of Life?

26.1 How Did the Living World Begin to Diversify?

You may think that you have little in common with unicellular prokaryotes. But multicellular eukaryotes like yourself actually share many attributes with Bacteria and Archaea. For example, all three of you:

- conduct glycolysis
- use DNA as the genetic material that encodes proteins
- produce those proteins by transcription and translation using a similar genetic code
- replicate DNA semiconservatively
- have plasma membranes and ribosomes in abundance

These features support the conclusion that all living organisms are related to one another. If life had multiple origins, there would be little reason to expect all organisms to use overwhelmingly similar genetic codes or to share structures as unique as ribosomes. Furthermore, similarities in DNA sequences of universal genes (such as those that encode the structural components of ribosomes) confirm the monophyly of life.

Despite the commonalities found across all three domains, major differences have evolved as well. Let’s first distinguish between Eukarya and the two prokaryotic domains. Note that “domain” is a subjective term used for the largest groups of life. There is no objective definition of a domain, any more than there is of a kingdom or a family.

The three domains differ in significant ways

Prokaryotic cells differ from eukaryotic cells in three important ways:

- *Prokaryotic cells lack a cytoskeleton and a nucleus, so they do not divide by mitosis.* Instead, after replicating their DNA (see Figure 11.2), prokaryotic cells divide by their own method, *binary fission*.
- *The organization of the genetic material differs.* The DNA of the prokaryotic cell is not organized within a membrane-enclosed nucleus. DNA molecules in prokaryotes (both bacteria and archaea) are often circular. Many (but not all) prokaryotes have only one main chromosome and are effectively haploid, although many have additional smaller DNA molecules, called *plasmids*, as well (see Section 12.6).

TABLE 26.1
The Three Domains of Life on Earth

CHARACTERISTIC	BACTERIA	DOMAIN ARCHAEA	EUKARYA
Membrane-enclosed nucleus	Absent	Absent	Present
Membrane-enclosed organelles	Absent	Absent	Present
Peptidoglycan in cell wall	Present	Absent	Absent
Membrane lipids	Ester-linked Unbranched	Ester-linked Branched	Ester-linked Unbranched
Ribosomes ^a	70S	70S	80S
Initiator tRNA	Formylmethionine	Methionine	Methionine
Operons	Yes	Yes	No
Plasmids	Yes	Yes	Rare
RNA polymerases	One	One ^b	Three
Ribosomes sensitive to chloramphenicol and streptomycin	Yes	No	No
Ribosomes sensitive to diphtheria toxin	No	Yes	Yes

^a70S ribosomes are smaller than 80S ribosomes.

^bArchaeal RNA polymerase is similar to eukaryotic polymerases.

- Prokaryotes have none of the membrane-enclosed cytoplasmic organelles—mitochondria, Golgi apparatus, and others—that are found in most eukaryotes. However, the cytoplasm of a prokaryotic cell may contain a variety of infoldings of the plasma membrane and photosynthetic membrane systems not found in eukaryotes.

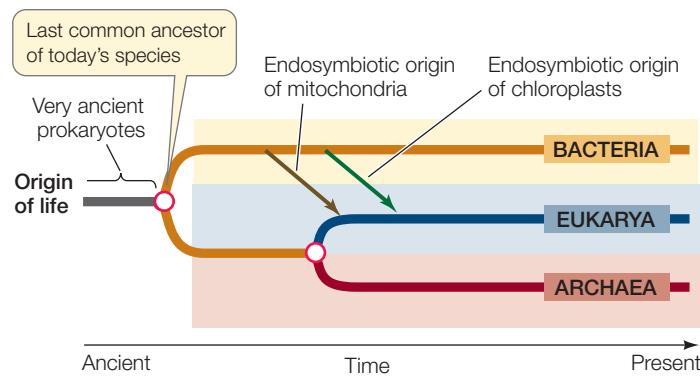
A glance at **Table 26.1** will show you that there are also major differences (most of which cannot be seen even under an electron microscope) between the two prokaryotic domains. In some ways archaea are more like eukaryotes; in other ways they are more like bacteria. (Note that we use lowercase when referring to the members of these domains and uppercase when referring to the domains themselves.) The structures of prokaryotic and eukaryotic cells are compared in Chapter 5. The basic unit of an *archaeon* (the term for a single archaeal organism) or *bacterium* (a single bacterial organism) is the prokaryotic cell. Each single-celled organism contains a full complement of genetic and protein-synthesizing systems, including DNA, RNA, and all the enzymes needed to transcribe and translate the genetic information into proteins. The prokaryotic cell also contains at least one system for generating the ATP it needs.

Genetic studies clearly indicate that all three domains had a single common ancestor. For a major portion of their genome, eukaryotes share a more recent common ancestor with Archaea than they do with Bacteria (**Figure 26.1**). However, the mitochondria of eukaryotes (as well as the chloroplasts of photosynthetic eukaryotes, such as plants) originated through the endosymbiosis of a bacterium, as described in Section 5.5. Some biologists prefer to view the origin of eukaryotes as a fusion of two equal

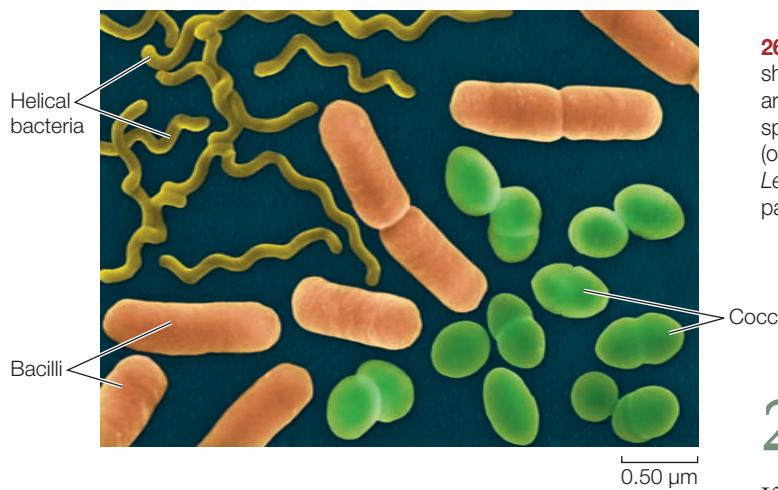
partners (one ancestor that was related to modern archaea, and another that was more closely related to modern bacteria). Others view the divergence of the early eukaryotes from the archaea as a separate and earlier event than the later endosymbiosis of the bacterium (the origin of mitochondria). In either case, some genes of eukaryotes are more closely related to those of archaea, while others are more closely related to those of bacteria. The tree of life therefore contains some merging of lineages, as well as the predominant diverging of lineages.

The common ancestor of all three domains had DNA as its genetic material, and its machinery for transcription and translation produced RNAs and proteins, respectively. This ancestor probably had a circular chromosome.

Three shapes are particularly common among the bacteria: spheres, rods, and curved or helical forms (**Figure 26.2**). A spherical bacterium is called a *coccus* (plural *cocci*). Coccii may live singly or may associate in two- or three-dimensional arrays



26.1 The Three Domains of the Living World All three domains share a common prokaryotic ancestor.



as chains, plates, blocks, or clusters of cells. A rod-shaped bacterium is called a *bacillus* (plural *bacilli*). The spiral form (like a corkscrew), or *helix* (plural *helices*), is the third main bacterial shape. Bacilli and helices may be single, form chains, or gather in regular clusters. Among the other bacterial shapes are long filaments and branched filaments.

Less is known about the shapes of archaea because many of these organisms have never been seen. Many archaea are known only from samples of DNA from the environment, as we describe in Section 26.4. However, the morphology of some species is known, including cocci, bacilli, and even triangular and square-shaped species; the latter grow on surfaces, arranged like sheets of postage stamps.

Archaea, Bacteria, and Eukarya are all products of billions of years of mutation, natural selection, and genetic drift, and they are all well adapted to present-day environments. None is “primitive.” Their last common ancestor probably lived 2 to 3 billion years ago. The earliest prokaryotic fossils date back at least 3.5 billion years, and they indicate that there was considerable diversity among the prokaryotes even during the earliest days of life.

26.1 RECAP

Bacteria and archaea are highly divergent from each other and are only distantly related on the tree of life. Eukaryotes received ancient evolutionary contributions from both of these prokaryotic lineages.

- What are the principal differences between the prokaryotes and the eukaryotes? **See pp. 537–538 and Table 26.1**
- Why don’t we group Bacteria and Archaea together in a single domain? **See p. 538 and Table 26.1**

The prokaryotes were alone on Earth for a very long time, adapting to new environments and to changes in existing environments. They have survived to this day, in massive numbers and incredible diversity, and they are found everywhere.

26.2 Bacterial Cell Shapes This composite, colorized micrograph shows the three common types of bacterial morphology. Spherical cells are called cocci; the acid-producing cocci shown here in green are a species of *Enterococcus* from the mammalian gut. The rod-shaped bacilli (orange) are represented by *Escherichia coli*, also a resident of the gut. *Leptospira interrogans* is a helical (spiral) bacterium and a human pathogen.

26.2 What Are Some Keys to the Success of Prokaryotes?

If success is measured by numbers of individuals, the prokaryotes are the most successful organisms on Earth. Individual bacteria and archaea in the oceans number more than 3×10^{28} . This stunning number is perhaps 100 million times as great as the number of stars in the visible universe. In fact, the bacteria living in your intestinal tract outnumber all the humans who have ever lived.

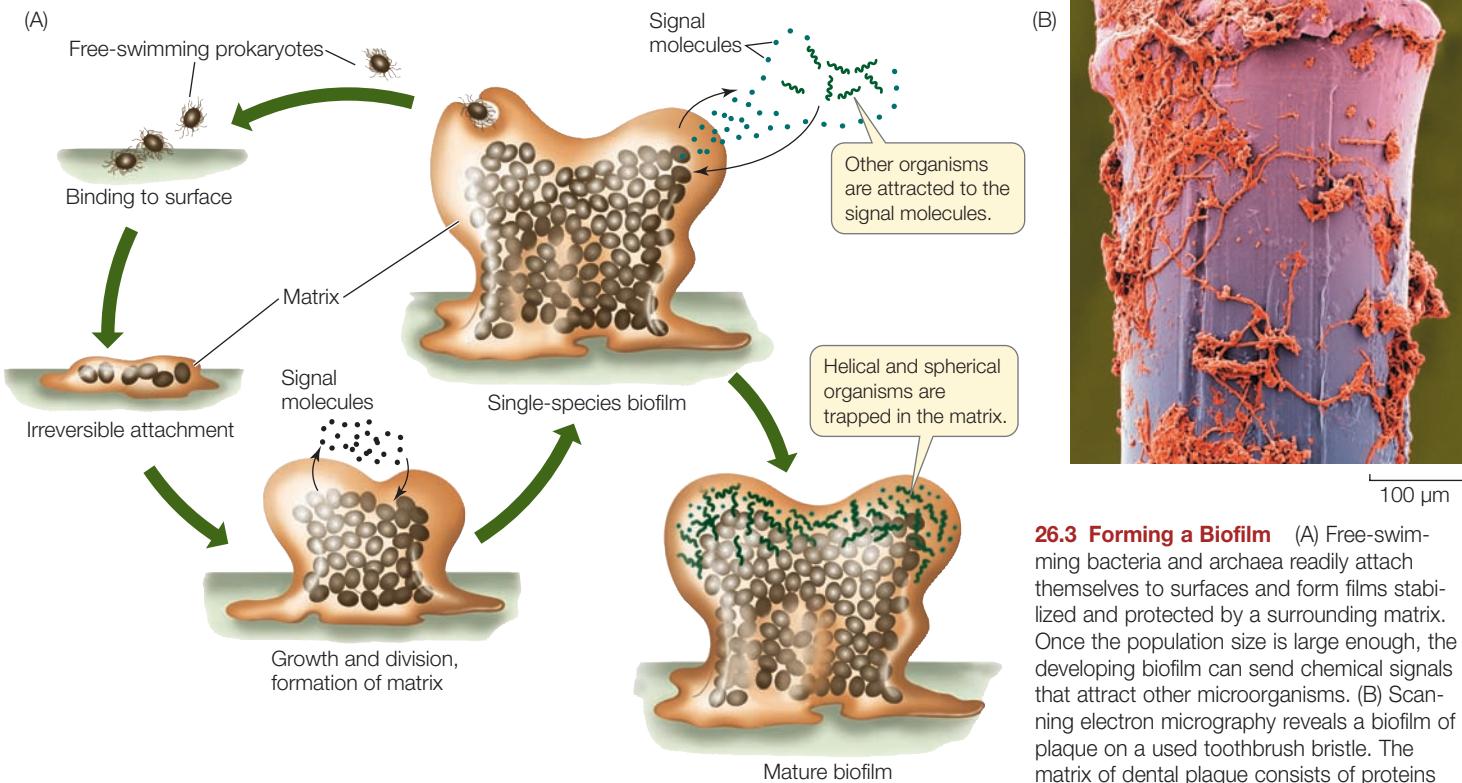
Prokaryotes are unicellular organisms, although many form multicellular colonies that contain many individual cells. These multicellular associations are not cases of true multicellular organisms, however, because each individual cell is fully viable and independent. These associations arise as cells adhere to one another after reproducing by binary fission. Associations in the form of chains are called **filaments**. Some filaments become enclosed in delicate tubular sheaths.

Prokaryotes generally form complex communities

Prokaryotic cells and their associations do not usually live in isolation. Rather, they live in communities of many different species of organisms, often including microscopic eukaryotes. (Microscopic organisms are often collectively referred to as *microbes*.) Some microbial communities form layers in sediments, and others form clumps a meter or more in diameter. While some microbial communities are harmful to humans, others provide important services. They help us digest our food, break down municipal waste, and recycle organic matter in the environment.

Many microbial communities tend to form dense **biofilms**. Upon contacting a solid surface, the cells secrete a gel-like sticky polysaccharide matrix that then traps other cells (Figure 26.3). Once this biofilm forms, it is difficult to kill the cells. Pathogenic (disease-causing) bacteria are difficult for the immune system—and modern medicine—to combat once they form a biofilm. For example, the film may be impermeable to antibiotics. Worse, some drugs stimulate the bacteria in a biofilm to lay down more matrix, making the film even more impermeable.

Biofilms form on contact lenses, on artificial joint replacements, and on just about any available surface. They foul metal pipes and cause corrosion, a major problem in steam-driven electricity generation plants. The stain on our teeth that we call dental plaque is also a biofilm. Fossil stromatolites—large, rocky structures made up of alternating layers of fossilized microbial biofilm and calcium carbonate—are the oldest remnants of early life on



26.3 Forming a Biofilm (A) Free-swimming bacteria and archaea readily attach themselves to surfaces and form films stabilized and protected by a surrounding matrix. Once the population size is large enough, the developing biofilm can send chemical signals that attract other microorganisms. (B) Scanning electron micrograph reveals a biofilm of plaque on a used toothbrush bristle. The matrix of dental plaque consists of proteins from both bacterial secretions and saliva.

Earth (see Figure 25.4). Stromatolites still form today in some parts of the world.

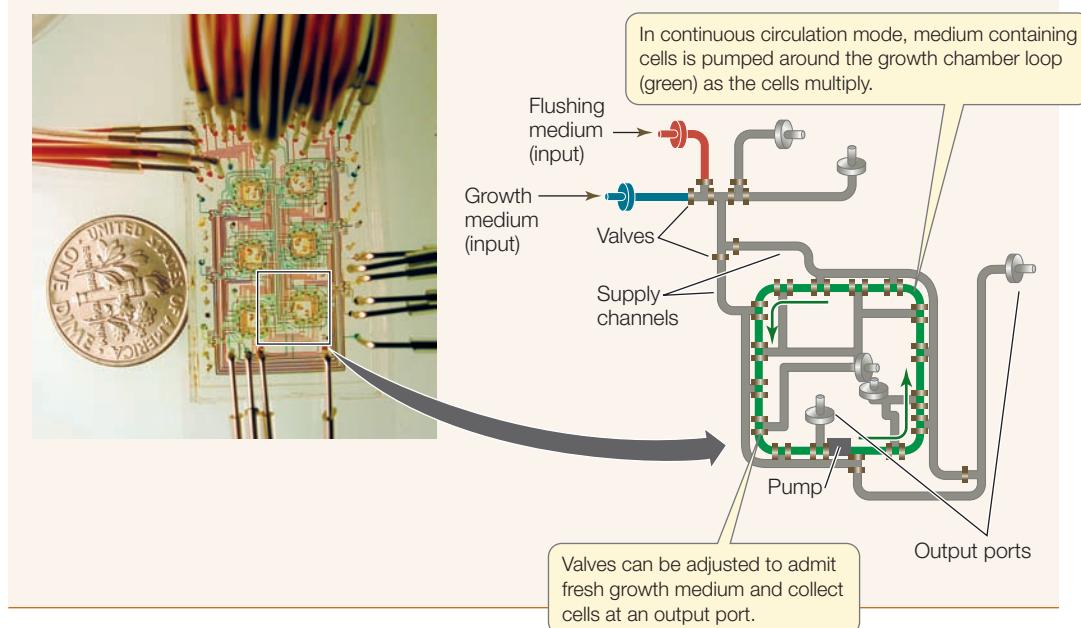
Biofilms are the subject of much current research. For example, some biologists are studying the chemical signals that bacteria in biofilms use to communicate with one another. By blocking the signals that lead to the production of the matrix polysaccharides, researchers may be able to prevent biofilms from forming.

A team of bioengineers and chemical engineers recently devised a sophisticated technique that enables them to monitor biofilm development in extremely small populations of bacteria, cell by cell. They developed a tiny chip housing six separate growth chambers, or “microchemostats” (Figure 26.4). The techniques of *microfluidics* use microscopic tubes and computer-controlled valves to direct fluid flow through complex “plumbing circuits” in the growth chambers.

TOOLS FOR INVESTIGATING LIFE

26.4 The Microchemostat

Using techniques from microfluidic engineering, biologists can monitor the dynamics of extremely small bacterial populations. The photograph shows six microchemostats on a chip. Each of the six is equipped with input ports for growth and flushing media, and a number of output ports (diagram). Tiny valves, controlled by a computer, direct flow. Samples are removed through the output ports and are analyzed to record changes in the bacterial population.



Prokaryotes have distinctive cell walls

Many prokaryotes have a thick and relatively stiff cell wall. It is quite different from the cell walls of land plants and algae, which contain cellulose and other polysaccharides, and from those of fungi, which contain chitin. The cell walls of almost all bacteria contain **peptidoglycan** (a cross-linked polymer of amino sugars), which produces a meshlike structure around the cell. Archaeal cell walls are of differing types, but most contain significant amounts of protein. One group of archaea has *pseudopeptidoglycan* in its cell wall; as you can probably guess from the prefix *pseudo*, pseudopeptidoglycan is similar to, but distinctly different from, the peptidoglycan of bacteria. The monomers making up pseudopeptidoglycan differ from and are differently linked than those of peptidoglycan. Peptidoglycan is a substance unique to bacteria; its absence from the

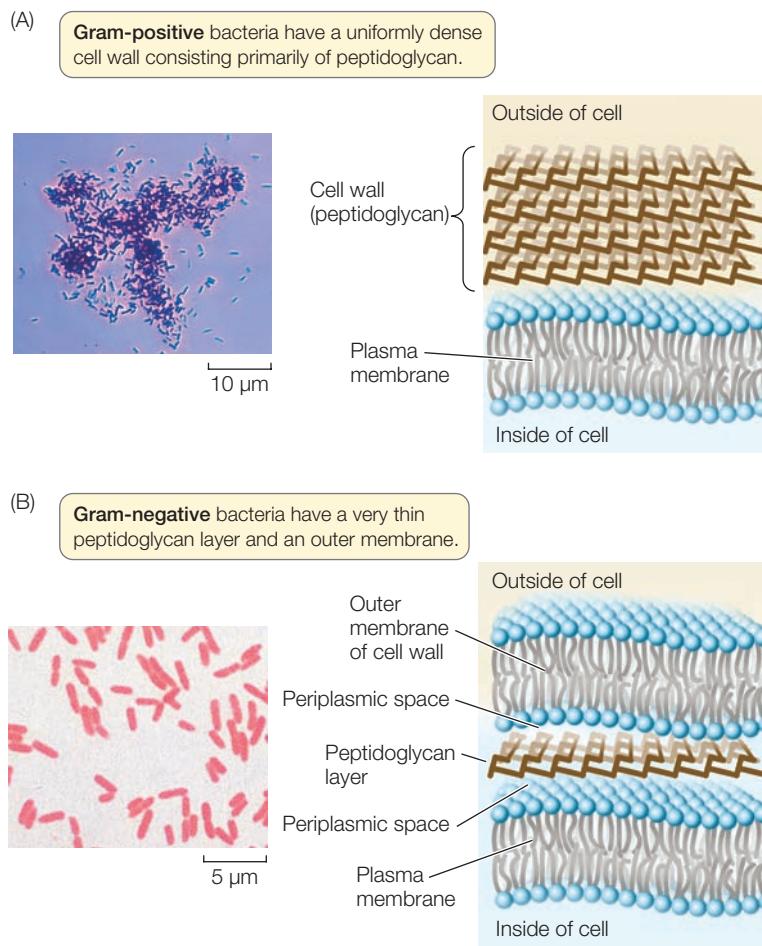
walls of archaea is a key difference between the two prokaryotic domains.

To appreciate the complexity of some bacterial cell walls, consider the reactions of bacteria to a simple staining process. A test called the **Gram stain** separates most types of bacteria into two distinct groups, Gram-positive and Gram-negative. A smear of cells on a microscope slide is soaked in a violet dye and treated with iodine; it is then washed with alcohol and counterstained with a red dye (safranine). **Gram-positive bacteria** retain the violet dye and appear blue to purple (**Figure 26.5A**). The alcohol washes the violet stain out of Gram-negative cells; these cells then pick up the safranine counterstain, so **Gram-negative bacteria** appear pink to red (**Figure 26.5B**).

For most bacteria, the Gram-staining results are determined by the chemical structure of the cell wall. A Gram-negative cell wall usually has a thin peptidoglycan layer, and outside the peptidoglycan layer the cell is surrounded by a second, outer membrane quite distinct in chemical makeup from the plasma membrane (see **Figure 26.5B**). Between the inner (plasma) and outer membranes of Gram-negative bacteria is a *periplasmic space*. This space contains proteins that are important in digesting some materials, transporting others, and detecting chemical gradients in the environment.

A Gram-positive cell wall usually has about five times as much peptidoglycan as a Gram-negative wall. This thick peptidoglycan layer is a meshwork that may serve some of the same purposes as the periplasmic space of the Gram-negative cell wall.

The consequences of the different features of prokaryotic cell walls are numerous and relate to the disease-causing characteristics of some bacteria. Indeed, the cell wall is a favorite target in medical combat against pathogenic bacteria because it has no counterpart in eukaryotic cells. Antibiotics such as penicillin and ampicillin, as well as other agents that specifically interfere with the synthesis of peptidoglycan-containing cell walls, tend to have little, if any, effect on the cells of humans and other eukaryotes.

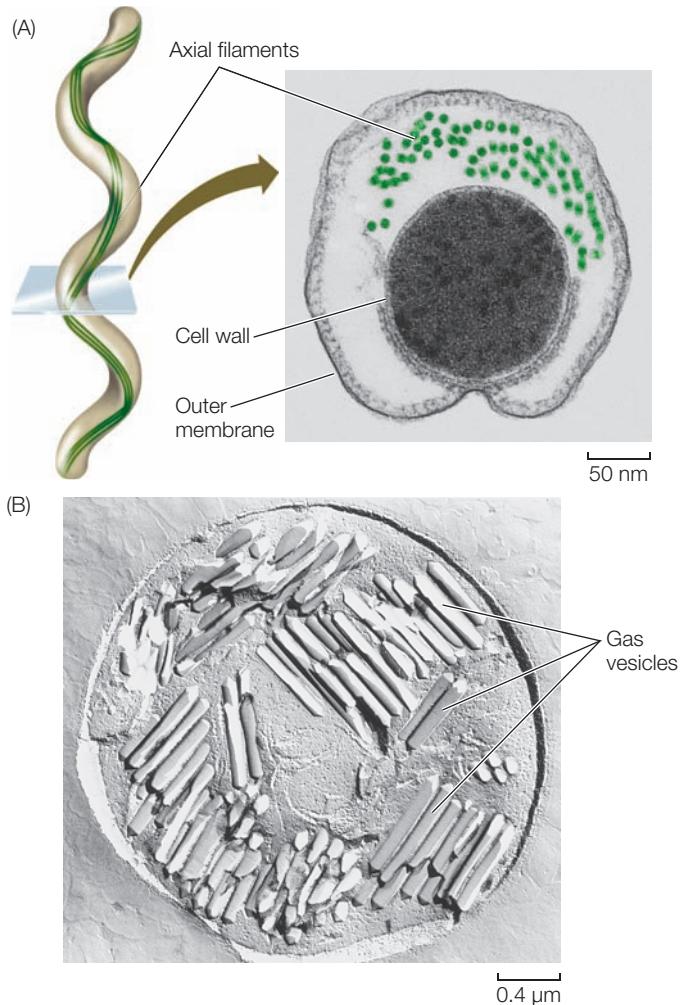


26.5 The Gram Stain and the Bacterial Cell Wall When treated with Gram stain, the cell walls of different bacteria react in one of two ways. (A) Gram-positive bacteria have a thick peptidoglycan cell wall that retains the violet dye and appears deep blue or purple. (B) Gram-negative bacteria have a thin peptidoglycan layer that does not retain the violet dye but picks up the counterstain and appears pink to red.

Prokaryotes have distinctive modes of locomotion

Although many prokaryotes cannot move, others are **motile**. These organisms move by one of several means. Some helical bacteria, called *spirochetes*, use a corkscrew-like motion made possible by modified flagella, called *axial filaments*, running along the axis of the cell beneath the outer membrane (**Figure 26.6A**). Many cyanobacteria and a few other groups of bacteria use various poorly understood gliding mechanisms, including rolling. Various aquatic prokaryotes, including some cyanobacteria, can move slowly up and down in the water by adjusting the amount of gas in *gas vesicles* (**Figure 26.6B**). By far the most common type of locomotion in prokaryotes, however, is that driven by flagella.

Prokaryotic **flagella** are slender filaments that extend singly or in tufts from one or both ends of the cell or are distributed all around it (**Figure 26.7**). A prokaryotic flagellum consists of a single fibril made of the protein fla-

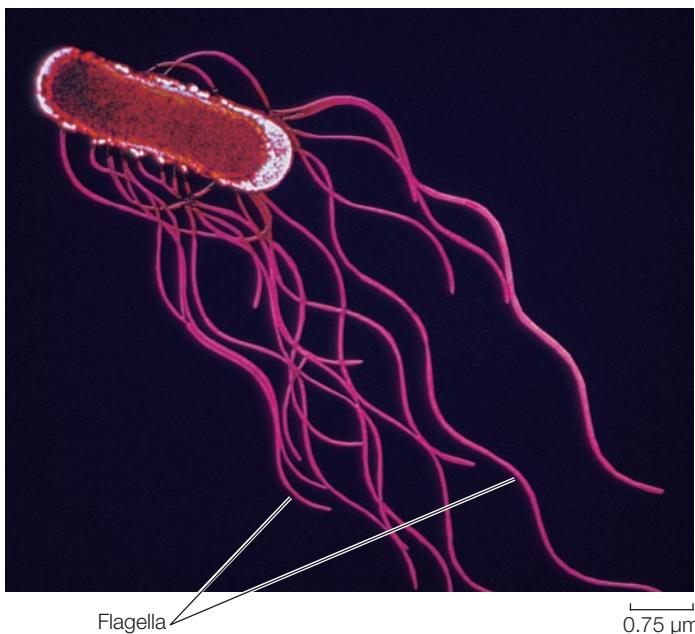


26.6 Structures Associated with Prokaryote Motility (A) A spirochete from the gut of a termite, seen in cross section, shows the axial filaments used to produce a corkscrew-like motion. (B) Gas vesicles in a cyanobacterium, visualized by the freeze-fracture technique.

gellin, projecting from the cell surface, plus a hook and basal body responsible for motion (see Figure 5.5). In contrast, the flagellum of eukaryotes is enclosed by the plasma membrane and usually contains a circle of nine pairs of microtubules surrounding two central microtubules, all containing the protein *tubulin*, along with many other associated proteins. The prokaryotic flagellum rotates about its base, much like a propeller, rather than beating in a whiplike manner, as a eukaryotic flagellum or cilium does.

Prokaryotes reproduce asexually, but genetic recombination can occur

Prokaryotes reproduce by binary fission, an asexual process (see Figure 11.2). Recall, however, that there are also processes—transformation, conjugation, and transduction—that allow the exchange of genetic information between some prokaryotes without reproduction occurring. So prokaryotes can exchange and recombine their DNA with other individuals (this is sex in



26.7 Some Prokaryotes Use Flagella for Locomotion Multiple flagella propel this *Salmonella* bacillus.

the genetic sense of the word), but this genetic exchange is not directly linked to reproduction as it is in most eukaryotes.

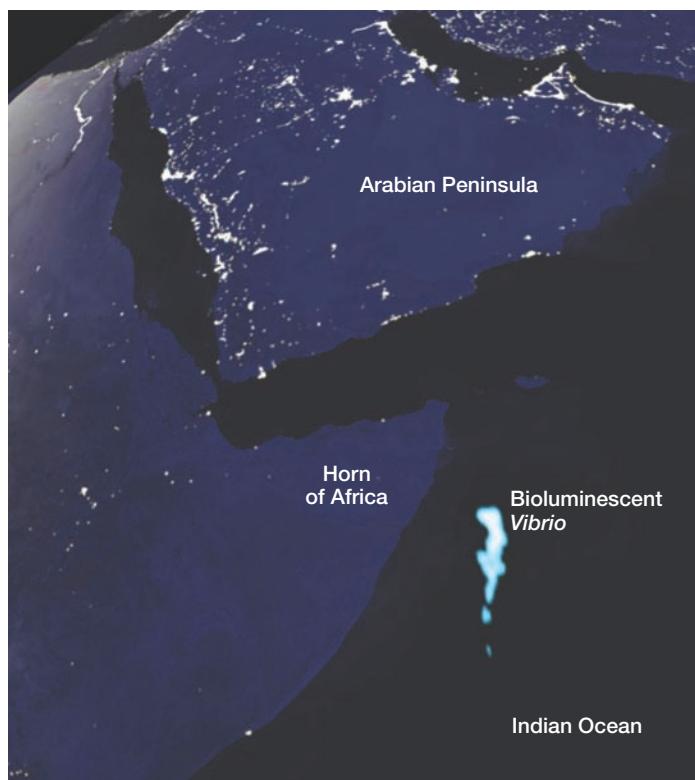
If conditions are favorable, some prokaryotes can multiply very rapidly. The shortest known prokaryote generation times are about 10 minutes, although these rapid rates of replication usually are not maintained for long. Under less optimal conditions, generation times often extend to many hours or even several days. Bacteria living deep in Earth's crust may suspend their growth for more than a century without dividing, then multiply for a few days before once again suspending growth.

Prokaryotes can communicate

Prokaryotes can send and receive signals from one another and from other organisms. One communication channel they employ is chemical. Another is physical, with light as the medium.

Bacteria release chemical substances that are sensed by other bacteria of the same species. They can announce their availability for conjugation, for example, by means of such signals. They can also monitor the density of their population. As the density of bacteria in a particular region increases, the concentration of a chemical signal builds up. When the bacteria sense that their population has become sufficiently dense, they can commence activities that smaller densities could not manage, such as forming a biofilm (see Figure 26.3). This density-sensing technique is called **quorum sensing**.

Like fireflies and many other organisms, some bacteria can emit light by a process called **bioluminescence**. A complex, enzyme-catalyzed reaction requiring ATP causes the emission of light but not heat. Often such bacteria luminesce only when a quorum has been sensed. The bioluminescent spots present in some deep-sea fishes are produced by colonies of biolumines-



26.8 Bioluminescent Bacteria Seen from Space In this satellite photo, legions of bioluminescent *Vibrio harveyi* form a glowing patch thousands of square kilometers in area in the Indian Ocean, off the Horn of Africa. Compare their blue glow with the white light of cities in eastern Africa and the Middle East.

cent bacteria. On land, some soil-dwelling bioluminescent bacteria produce eerily glowing patches of ground at night.

How is bioluminescence useful to a prokaryote? One fairly well understood case is that of some bacteria of the genus *Vibrio*. These bacteria can live freely, but they truly thrive inside the guts of fish. Inside the fish, they may attach to food particles and then can be expelled as waste along with particulate matter. Reproducing on the particles, a bacteria population increases until a glowing particle attracts another fish, which ingests the bacteria along with the particle—giving the bacteria a new home and food source for a while. In this case, *Vibrio* are both communicating with another species and enhancing their own nutritional status. In the Indian Ocean off the eastern coast of Africa, *Vibrio* sometimes concentrate over such a large area (several thousand square kilometers) that their bioluminescence is visible from space (**Figure 26.8**).

Prokaryotes have amazingly diverse metabolic pathways

Bacteria and archaea outdo the eukaryotes in terms of metabolic diversity. Although much more diverse in size and shape, eukaryotes draw on fewer metabolic mechanisms for their energy needs. In fact, much of the eukaryotes' energy metabolism is carried out in organelles—mitochondria and chloro-

plasts—that are endosymbiotic descendants of bacteria, as described in Section 5.5.

The long evolutionary history of bacteria and archaea, during which they have had time to explore a wide variety of habitats, has led to the extraordinary diversity of their metabolic “lifestyles”—their use or nonuse of oxygen, their energy sources, their sources of carbon atoms, and the materials they release as waste products.

ANAEROBIC VERSUS AEROBIC METABOLISM Some prokaryotes can live only by anaerobic metabolism because molecular oxygen is poisonous to them. These oxygen-sensitive organisms are called **obligate anaerobes**. Other prokaryotes can shift their metabolism between anaerobic and aerobic modes (see Chapter 9) and thus are called **facultative anaerobes**. Many facultative anaerobes alternate between anaerobic metabolism (such as fermentation) and cellular respiration as conditions dictate. **Aerotolerant anaerobes** cannot conduct cellular respiration but are not damaged by oxygen when it is present. By definition, an anaerobe does not use oxygen as an electron acceptor for its respiration.

At the other extreme from the obligate anaerobes, some prokaryotes are **obligate aerobes**, unable to survive for extended periods in the *absence* of oxygen. They require oxygen for cellular respiration.

NUTRITIONAL CATEGORIES All living organisms face the same nutritional challenges: they must synthesize energy-rich compounds such as ATP to power their life-sustaining metabolic reactions, and they must obtain carbon atoms to build their own organic molecules. Biologists recognize four broad nutritional categories of organisms: photoautotrophs, photoheterotrophs, chemolithotrophs, and chemoheterotrophs. Prokaryotes are represented in all four groups (**Table 26.2**).

Photoautotrophs perform photosynthesis. They use light as their energy source and carbon dioxide (CO₂) as their carbon source. Like green plants and other photosynthetic eukaryotes, the cyanobacteria, a group of photoautotrophic bacteria, use chlorophyll *a* as their key photosynthetic pigment and produce

TABLE 26.2		
How Organisms Obtain Their Energy and Carbon		
NUTRITIONAL CATEGORY	ENERGY SOURCE	CARBON SOURCE
Photoautotrophs (found in all three domains)	Light	Carbon dioxide
Photoheterotrophs (some bacteria)	Light	Organic compounds
Chemolithotrophs (some bacteria, many archaea)	Inorganic substances	Carbon dioxide
Chemoheterotrophs (found in all three domains)	Organic compounds	Organic compounds

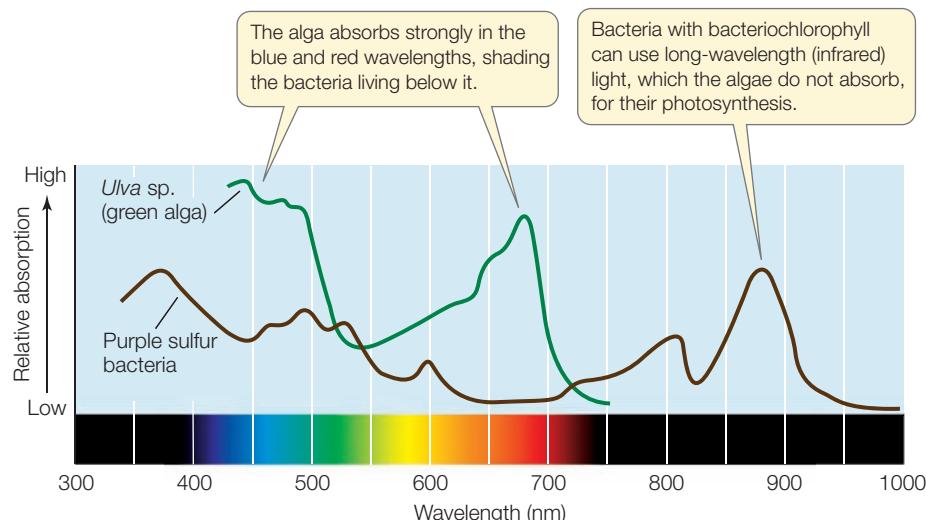
oxygen gas (O_2) as a by-product of noncyclic electron transport (see Section 10.1).

There are other photosynthetic groups among the bacteria, but these use *bacteriochlorophyll* as their key photosynthetic pigment, and they do not release O_2 . Indeed, some of these photosynthesizers produce particles of pure sulfur, because hydrogen sulfide (H_2S) rather than H_2O is their electron donor for photophosphorylation (see Section 10.2). *Bacteriochlorophyll* molecules absorb light of longer wavelengths than the chlorophyll molecules used by all other photosynthesizing organisms. As a result, bacteria using this pigment can grow in water under fairly dense layers of algae, using light of wavelengths that are not absorbed by the algae (Figure 26.9).

Photoheterotrophs use light as their energy source but must obtain their carbon atoms from organic compounds made by other organisms. Their “food” consists of organic compounds such as carbohydrates, fatty acids, and alcohols. For example, compounds released from plant roots (as in rice paddies) or from decomposing photosynthetic bacteria in hot springs are taken up by photoheterotrophs and metabolized to form building blocks for other compounds; sunlight provides the necessary ATP through photophosphorylation. The purple nonsulfur bacteria, among others, are photoheterotrophs.

Chemolithotrophs (also called chemoautotrophs) obtain their energy by oxidizing inorganic substances, and they use some of that energy to fix CO_2 . Some chemolithotrophs use reactions identical to those of the typical photosynthetic cycle, but others use alternative pathways to fix CO_2 . Some bacteria oxidize ammonia or nitrite ions to form nitrate ions. Others oxidize hydrogen gas, hydrogen sulfide, sulfur, and other materials. Many archaea are chemolithotrophs.

Deep-sea hydrothermal vent ecosystems are dependent on chemolithotrophic prokaryotes that are incorporated into large communities of crabs, mollusks, and giant worms, all living at a depth of 2,500 meters—below any hint of sunlight. These bacteria obtain energy by oxidizing hydrogen sulfide and other substances released in the near-boiling water flowing from volcanic vents in the ocean floor.

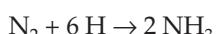


Finally, **chemoheterotrophs** obtain both energy and carbon atoms from one or more complex organic compounds that have been synthesized by other organisms. Most known bacteria and archaea are chemoheterotrophs—as are all animals and fungi and many protists.

NITROGEN AND SULFUR METABOLISM Key metabolic reactions in many prokaryotes involve nitrogen or sulfur. For example, some bacteria carry out respiratory electron transport without using oxygen as an electron acceptor. These organisms use oxidized inorganic ions such as nitrate, nitrite, or sulfate as electron acceptors. Examples include the **denitrifiers**, bacteria that release nitrogen to the atmosphere as nitrogen gas (N_2). These normally aerobic bacteria, mostly species of the genera *Bacillus* and *Pseudomonas*, use nitrate (NO_3^-) as an electron acceptor in place of oxygen if they are kept under anaerobic conditions:



Nitrogen fixers convert atmospheric nitrogen gas into a chemical form (ammonia) usable by the nitrogen fixers themselves as well as by other organisms, especially land plants:



All organisms require nitrogen in order to build proteins, nucleic acids, and other important compounds. Nitrogen fixation is thus vital to life as we know it. This all-important biochemical process is carried out by a wide variety of archaea and bacteria (including cyanobacteria) but by no other organisms, so we depend on these prokaryotes for our very existence. We describe the details of nitrogen fixation in Chapter 36.

Ammonia is oxidized to nitrate in soil and in seawater by chemolithotrophic bacteria called **nitrifiers**. Bacteria of two genera, *Nitrosomonas* and *Nitrosococcus*, convert ammonia to nitrite ions (NO_2^-), and *Nitrobacter* oxidizes nitrite to nitrate (NO_3^-).

What do the nitrifiers get out of these reactions? Their metabolism is powered by the energy released by the oxidation of ammonia or nitrite. For example, by passing the electrons from nitrite through an electron transport chain (see Section 9.3),

Nitrobacter can make ATP, and using some of this ATP, can also make NADH. With this ATP and NADH, the bacterium can convert CO_2 and H_2O to glucose.

26.9 Bacteriochlorophyll Absorbs Long-Wavelength Light

The chlorophyll in *Ulva*, a green alga, absorbs no light of wavelengths longer than 750 nm. Purple sulfur bacteria, which contain bacteriochlorophyll, can conduct photosynthesis using longer wavelengths.

26.2 RECAP

Prokaryotes have established themselves everywhere on Earth. They may form communities called biofilms that coat materials with a gel-like matrix. Prokaryotes have distinctive cell walls and modes of locomotion, communication, reproduction, and nutrition.

- How do biofilms form and why are they of special interest to researchers? **See pp. 539–540 and Figure 26.3**
- Describe bacterial cell wall architecture. **See p. 541 and Figure 26.5**
- How are the four nutritional categories of prokaryotes distinguished? **See pp. 543–544 and Table 26.2**
- Explain why nitrogen metabolism in the prokaryotes is vital to other organisms. **See p. 544**

We noted earlier that only recently have scientists appreciated the huge distinctions between Bacteria and Archaea. How do researchers approach the classification of organisms they can't even see?

26.3 How Can We Resolve Prokaryote Phylogeny?

As detailed in Chapter 22, classification schemes serve three primary purposes: to identify organisms, to reveal evolutionary relationships, and to provide universal names. Classifying bacteria and archaea is of particular importance to humans because scientists and medical technologists must be able to identify bacteria quickly and accurately; when the bacteria are pathogenic, lives may depend on it. In addition, many emerging biotechnologies (see Chapter 18) depend on a thorough knowledge of prokaryote biochemistry, and understanding an organism's phylogeny allows biologists to make predictions about the distribution of biochemical processes across the wide diversity of prokaryotes.

The small size of prokaryotes has hindered our study of their phylogeny

Until about 300 years ago, nobody had even seen an individual prokaryote; these organisms remained invisible to humans until the invention of the first simple microscope. Prokaryotes are so small that even the best light microscopes don't reveal much about them. It took the advanced microscopic equipment and techniques of the twentieth century (see Figure 5.3) to open up the microbial world.

Until recently, taxonomists based prokaryote classification on observable phenotypic characters such as shape, color, motility, nutritional requirements, antibiotic sensitivity, and reaction to the Gram stain. When biologists learned how to grow bacteria in pure culture on nutrient media, they learned a great deal about the genetics, nutrition, and metabolism of those species that could be cultured. However, these species represent prob-

ably less than 1 percent of living prokaryote species. Furthermore, this work provided little insight into how prokaryotic organisms evolved—a question of great interest to microbiologists and evolutionary biologists. Only recently have systematists developed the appropriate tools to produce classification schemes that make sense in evolutionary terms.

The nucleotide sequences of prokaryotes reveal their evolutionary relationships

Analyses of nucleotide sequences of ribosomal RNA (rRNA) genes provided the first comprehensive evidence of evolutionary relationships among prokaryotes. For several reasons, rRNA is particularly useful for evolutionary studies of living organisms:

- rRNA is evolutionarily ancient, as it was found in the common ancestor of life.
- No free-living organism lacks rRNA, so rRNA genes can be compared throughout the tree of life.
- rRNA plays a critical role in translation in all organisms, so *lateral transfer* of rRNA genes among distantly related species is unlikely.
- rRNA has evolved slowly enough that gene sequences can be aligned and analyzed among even distantly related species.

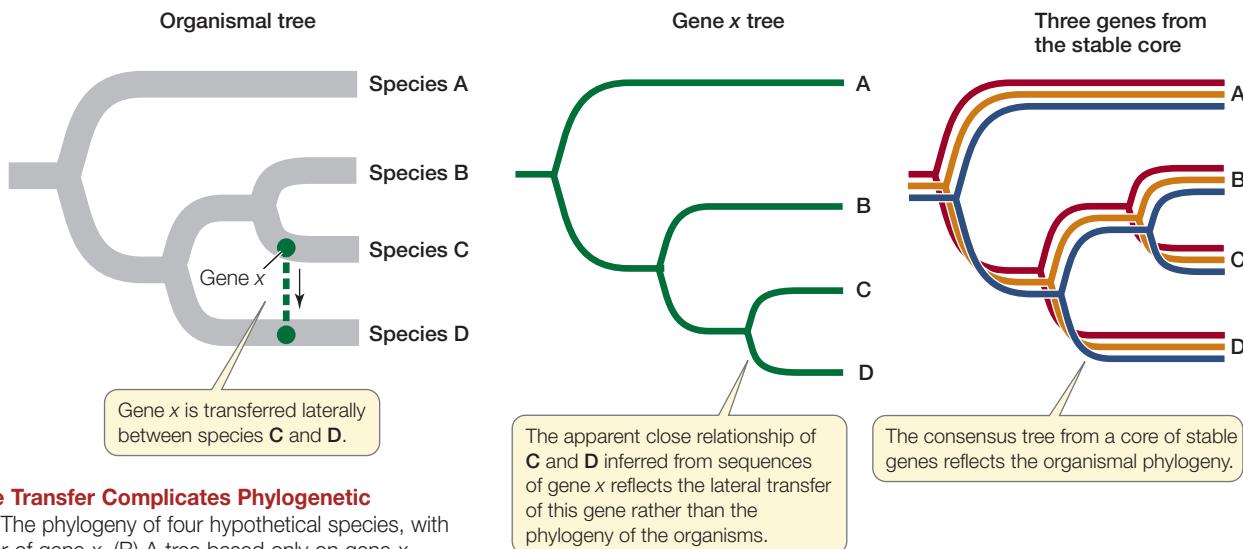
Comparisons of rRNA genes from a great many organisms have revealed the probable phylogenetic relationships from throughout the tree of life. Databases such as GenBank contain rRNA gene sequences from hundreds of thousands of species—more than any other gene sequences.

Although these data are helpful, it is clear that even distantly related prokaryotes sometimes exchange genetic material. In some groups of prokaryotes, analyses of multiple gene sequences have suggested several different phylogenetic patterns. How could such differences among different gene sequences arise?

Lateral gene transfer can lead to discordant gene trees

As noted earlier, prokaryotes reproduce by binary fission. If we could follow these divisions back through evolutionary time, we would be tracing the path of the complete tree of life for bacteria and archaea. This underlying tree of relationships, represented in highly abbreviated form in Appendix A, is called the *organismal* (or *species*) tree. Because whole genomes are replicated during asexual binary fission divisions, we expect phylogenetic trees constructed from most gene sequences to reflect these same relationships (see Chapter 22).

From early in evolution to the present day, however, some genes have been moving “sideways” from one prokaryotic species to another, a phenomenon known as **lateral gene transfer**. Mechanisms of lateral gene transfer include transfer by plasmids and viruses and uptake of DNA from the environment by transformation. Lateral gene transfers are well documented, especially among closely related species; some have been documented even across the three primary domains of life.



26.10 Lateral Gene Transfer Complicates Phylogenetic Relationships

(A) The phylogeny of four hypothetical species, with a lateral gene transfer of gene x. (B) A tree based only on gene x shows the phylogeny of the laterally transferred gene, rather than the organismal phylogeny. (C) In many cases, a “stable core” of prokaryotic genes can be used to reconstruct the organismal phylogeny of prokaryotes.

Consider, for example, the genome of *Thermotoga maritima*, a bacterium that can survive extremely high temperatures. In comparing the 1,869 gene sequences of *T. maritima* against sequences for the same proteins in other species, investigators found that some of this bacterium’s genes have their closest relationships not with those of other bacterial species, but with the genes of archaeal species that live in similar environments.

When genes involved in lateral transfer events are sequenced and analyzed phylogenetically, the resulting individual gene trees will not match the organismal phylogeny in every respect (**Figure 26.10**). Individual gene trees will vary because the history of lateral gene transfer events is different for each gene. Biologists reconstruct the underlying organismal phylogeny by comparing multiple genes (to produce a *consensus tree*), or by concentrating on genes that are unlikely to be involved in lateral gene transfer events. For example, genes that are involved in fundamental cell processes (such as the rRNA genes discussed above) are unlikely to be replaced by the same genes from other species, since functional, locally adapted copies of these genes are already present.

What kinds of genes are most likely to be involved in lateral gene transfer? Genes that result in a new, adaptive function that will convey higher fitness to a recipient species are most likely to be transferred repeatedly among species. For example, genes that produce antibiotic resistance are often transferred on plasmids among many bacterial species, especially under the strong selective conditions of antibiotic medication by humans. This selection for antibiotic resistance is why informed physicians are now more careful in prescribing antibiotics. Improper or frequent use of antibiotics can lead to selection for resistant strains of bacteria, which are then much harder to treat effectively.

It is debatable whether lateral gene transfer has seriously complicated our attempts to resolve the tree of prokaryotic life.

Recent work suggests that it has not—while it complicates studies in some individual species, it need not present problems at higher levels. It is now possible to make nucleotide sequence comparisons involving entire genomes, and these studies are revealing a *stable core* of crucial genes that are uncomplicated by lateral gene transfer. Gene trees based on this stable core more accurately reveal relationships of the organismal phylogeny (see Figure 26.10). The problem remains, however, that only a very small proportion of the prokaryotic world has been described and studied.

The great majority of prokaryote species have never been studied

Most prokaryotes have defied all attempts to grow them in pure culture, causing biologists to wonder how many species, and possibly even important clades, we might be missing. A window onto this problem was opened with the introduction of a new way to look at nucleic acid sequences. Unable to work with the whole genome of a single species, biologists instead examine sequences in individual genes collected from a random sample of the environment.

Norman Pace of the University of Colorado isolated individual rRNA gene sequences from extracts of environmental samples such as soil and seawater. Comparing such sequences with previously known ones revealed an extraordinary number of new sequences, implying that they came from previously unrecognized species. Biologists have described only about 10,000 species of bacteria and only a few hundred species of archaea (see Figure 1.10). The results of Pace’s and similar studies suggest that there may be millions, perhaps hundreds of millions, of prokaryote species on Earth. Other biologists put the estimate much lower, and argue that the high dispersal ability of many bacterial species greatly reduces local endemism (geographically restricted species). Only the magnitude of these estimates differ, however; all sides agree that we have just begun to uncover Earth’s bacterial and archaeal diversity.

26.3 RECAP

The study of prokaryote phylogeny and diversity has been inhibited by the organisms' small size, our inability to grow some of them in pure culture, and lateral gene transfer. However, nucleotide sequences of essential genes are providing a much clearer picture of bacterial and archaeal evolutionary relationships.

- How did biologists classify bacteria before it became possible to determine nucleotide sequences? **See p. 545**
- Explain why nucleotide sequences of rRNA genes are useful for evolutionary studies. **See p. 545**
- How does lateral gene transfer complicate evolutionary studies? **See p. 545–546 and Figure 26.10**

With the advent of sequencing techniques, biologists have made rapid progress in understanding the phylogeny of prokaryotes. In the next section, we identify the characteristics and life history of the major groups.

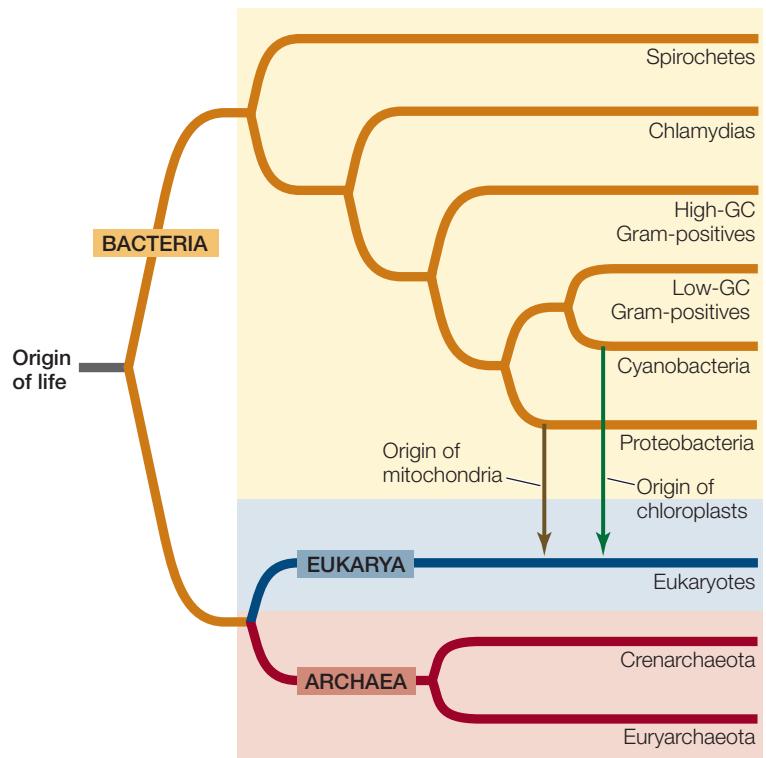
26.4 What Are the Major Known Groups of Prokaryotes?

Here we use a widely accepted classification scheme that has considerable support from nucleotide sequence data. More than a dozen major clades have been proposed under this scheme, just a few of which we discuss here. We pay the closest attention to six groups that have received the most study: the spirochetes, chlamydias, high-GC Gram-positives, cyanobacteria, low-GC Gram-positives, and proteobacteria (Figure 26.11). First, however, a few words about the origins of the prokaryotes are in order.

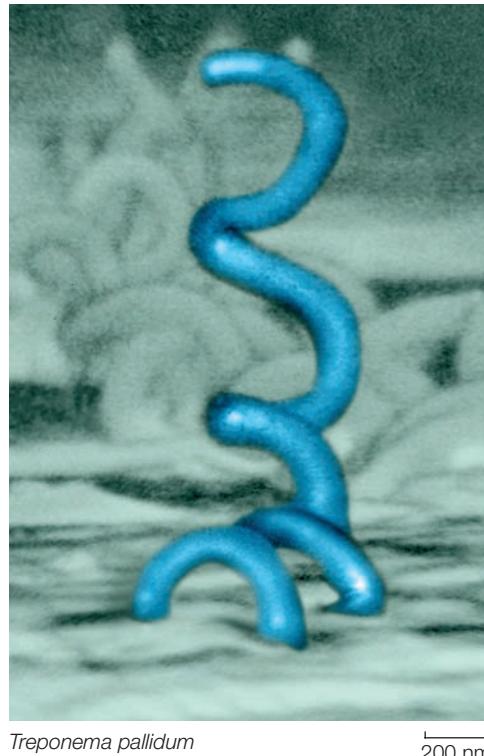
Several of the earliest branching lineages of bacteria and archaea are **thermophiles** (Greek, "heat-lovers"). This observation is in line with the hypothesis that the first living organisms were thermophiles, given that most environments on early Earth were much hotter than those of today. While additional evidence continues to support this hypothesis, some researchers believe that the various thermophilic groups evolved more recently than did the lineages leading to the spirochetes and chlamydias.

Spirochetes move by means of axial filaments

Spirochetes are Gram-negative, motile, chemoheterotrophic bacteria characterized by unique structures called axial filaments, which are modified flagella running through the periplasmic space (see Figure 26.6A). The cell body is a long cylinder coiled into a helix (Figure 26.12). The axial filaments begin at either end of the cell and overlap in the middle. Protein motors connect the axial filaments to the cell wall, enabling rotation of these structures as they do in other prokaryotic flagella. Many spirochetes live in humans as parasites; a few are pathogens, including those that cause syphilis and Lyme disease. Others live free in mud or water.



26.11 Two Domains: A Brief Overview This abridged summary classification of the domains Bacteria and Archaea shows their relationships to each other and to Eukarya. The relationships among the many clades of bacteria, not all of which are listed here, are incompletely resolved at this time.



26.12 A Spirochete This corkscrew-shaped bacterium causes syphilis in humans.

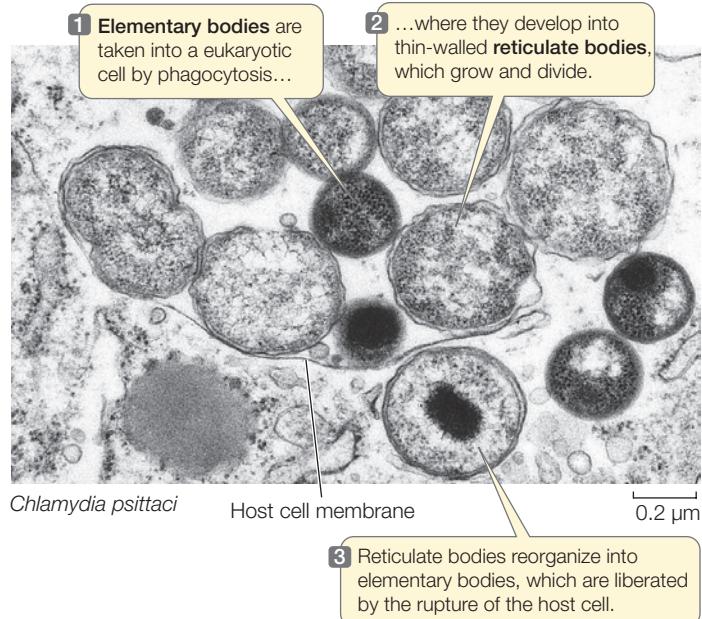
Chlamydias are extremely small parasites

Chlamydias are among the smallest bacteria ($0.2\text{--}1.5\ \mu\text{m}$ in diameter). They can live only as parasites in the cells of other organisms. It was once believed that this obligate parasitism resulted from an inability of chlamydias to produce ATP—that chlamydias were “energy parasites.” However, genome sequencing from the end of the twentieth century indicates that chlamydias have the genetic capability to produce at least some ATP. They can augment this capacity by using an enzyme called a translocase, which allows them to take up ATP from the cytoplasm of their host in exchange for ADP from their own cells.

These tiny, Gram-negative cocci are unique prokaryotes because of their complex life cycle, which involves two different forms of cells, *elementary bodies* and *reticulate bodies* (Figure 26.13). In humans, various strains of chlamydias cause eye infections (especially trachoma), sexually transmitted diseases, and some forms of pneumonia.

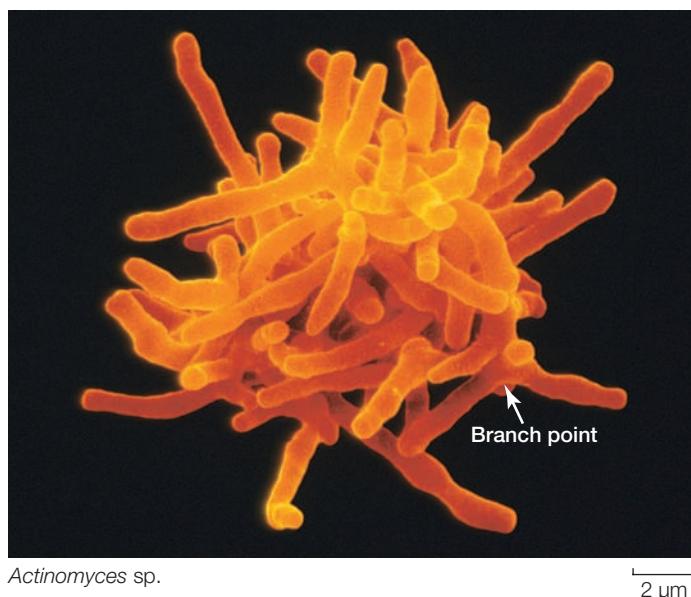
Some high-GC Gram-positives are valuable sources of antibiotics

High-GC Gram-positives, also known as *actinobacteria*, derive their name from the relatively high ratio of G-C to A-T nucleotide base pairs in their DNA. These bacteria develop an elaborately branched system of filaments (Figure 26.14) and can resemble the filamentous growth habit of fungi, albeit at a reduced scale. Some high-GC Gram-positives reproduce by forming chains of spores at the tips of the filaments. In species that do not form spores, the branched, filamentous growth ceases, and the structure breaks up into typical cocci or bacilli, which then reproduce by binary fission.



26.13 Chlamydias Change Form during their Life Cycle

Elementary bodies and reticulate bodies are the two major phases of the chlamydia life cycle.



26.14 Filaments of a High-GC Gram-Positive The branching filaments seen in this scanning electron micrograph are typical of this medically important bacterial group.

The high-GC Gram-positives include several medically important bacteria. *Mycobacterium tuberculosis* causes tuberculosis, which kills 3 million people each year. Genetic data suggest that this bacterium arose 3 million years ago in East Africa, making it the oldest known human bacterial affliction. *Streptomyces* produce streptomycin as well as hundreds of other antibiotics. We derive most of our antibiotics from members of the high-GC Gram-positives.

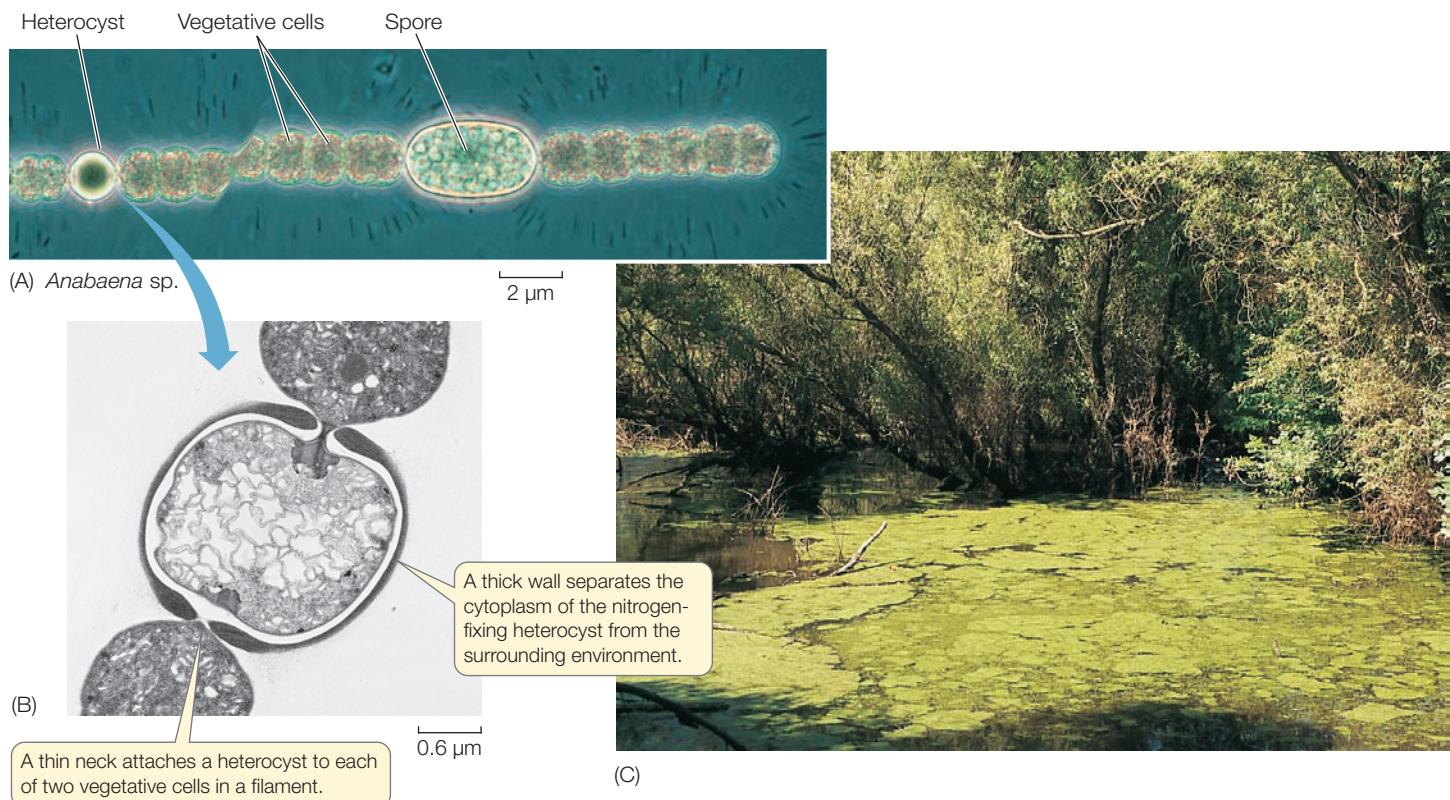
Cyanobacteria are important photoautotrophs

Cyanobacteria, sometimes called *blue-green bacteria* because of their pigmentation, are photoautotrophs that require only water, nitrogen gas, oxygen, a few mineral elements, light, and carbon dioxide to survive. They use chlorophyll *a* for photosynthesis and release oxygen gas; many species also fix nitrogen. Their photosynthesis was the basis of the “oxygen revolution” that transformed Earth’s atmosphere (see Section 25.3).

Cyanobacteria carry out the same type of photosynthesis that is characteristic of eukaryotic photosynthesizers. They contain elaborate and highly organized internal membrane systems called *photosynthetic lamellae*. The chloroplasts of photosynthetic eukaryotes are derived from an endosymbiotic cyanobacterium.

Cyanobacteria may live free as single cells or associate in colonies. Depending on the species and on growth conditions, colonies may range from flat sheets one cell thick to filaments to spherical balls of cells.

Some filamentous colonies of cyanobacteria differentiate into three cell types: vegetative cells, spores, and heterocysts (Figure 26.15). **Vegetative cells** photosynthesize, **spores** are resting stages that can survive harsh environmental conditions and eventually develop into new filaments, and **heterocysts** are cells specialized for nitrogen fixation. All of the known cyanobacteria with heterocysts fix nitrogen. Heterocysts also have a role in reproduction: when filaments break apart to reproduce, the heterocyst may serve as a breaking point.



26.15 Cyanobacteria (A) *Anabaena* is a genus of cyanobacteria that form filamentous colonies containing three cell types. (B) Heterocysts are specialized for nitrogen fixation and serve as a breaking point when filaments reproduce. (C) Cyanobacteria appear in enormous numbers in some environments. This California pond has experienced eutrophication: phosphorus and other nutrients generated by human activity have accumulated, feeding an immense green mat (commonly referred to as “pond scum”) that is made up of several species of free-living cyanobacteria.

The low-GC Gram-positives include the smallest cellular organisms

As their name suggests, the **low-GC Gram-positives** have a lower ratio of G-C to A-T nucleotide base pairs than do the high-GC Gram-positives. Some of the low-GC Gram-positives are in fact Gram-negative, and some have no cell wall at all. Despite these differences among the various species, phylogenetic analyses of DNA sequences support the monophyly of this clade.

Some low-GC Gram-positives can produce heat-resistant resting structures called **endospores** (Figure 26.16). When a key nutrient such as nitrogen or carbon becomes scarce, the bacterium replicates its DNA and encapsulates one copy, along with some of its cytoplasm, in a tough cell wall heavily thickened with peptidoglycan and surrounded by a spore coat. The parent cell then breaks down, releasing the endospore. Endospore production is not a reproductive process; the endospore merely replaces the parent cell. The endospore, however, can survive harsh environmental conditions that would kill the parent cell, such as high or low temperatures or drought, because it is *dormant*—its normal activity is suspended. Later, if it encounters favorable conditions, the endospore becomes metabolically active and divides, forming new cells that are like the parent cells.

Some endospores can be reactivated after more than 1,000 years of dormancy. There are even credible claims of reactivation of *Bacillus* endospores after millions of years. Members of this endospore-forming group of low-GC Gram-positives include the many species of *Clostridium* and *Bacillus*.

Dormant endospores of *Bacillus anthracis* are the source of an *exotoxin* (see page 555) that causes anthrax. The spores germinate when they sense specific molecules (macrophages; see Chapter 42) in the cytoplasm of mammalian blood cells. However, endospores of other, nonpathogenic *Bacillus* species do not germinate.



26.16 A Structure for Waiting Out Bad Times This low-GC Gram-positive bacterium, which can cause severe colitis in humans, produces endospores as resistant resting structures.



26.17 Low-GC Gram-Positives “Grape clusters” are the usual arrangement of staphylococci.

nate in this environment. *B. anthracis* has been used as a bioterrorism agent because large quantities of its endospores are relatively easy to transport and spread among human populations, where they may be inhaled or ingested.

The genus *Staphylococcus*—the **staphylococci**—includes low-GC Gram-positives that are abundant on the human body surface; they are responsible for boils and many other skin problems (Figure 26.17). *Staphylococcus aureus* is the best-known human pathogen in this genus; it is found in 20 to 40 percent of normal adults (and in 50 to 70 percent of hospitalized adults). In addition to skin diseases, it can cause respiratory, intestinal, and wound infections.

Another interesting group of low-GC Gram-positives, the **mycoplasmas**, lack cell walls, although some have a stiffening material outside the plasma membrane. The mycoplasmas include the smallest cellular creatures known (Figure 26.18)—even smaller than chlamydias. The smallest mycoplasmas capable of multiplying by fission have a diameter of about 0.2 μm. They are small in another crucial sense as well: they have less than half as much DNA as most other prokaryotes. It has been speculated that the amount of DNA in a mycoplasma, which codes for fewer than 500 proteins, may be the minimum amount required to encode the essential properties of a living cell.

The proteobacteria are a large and diverse group

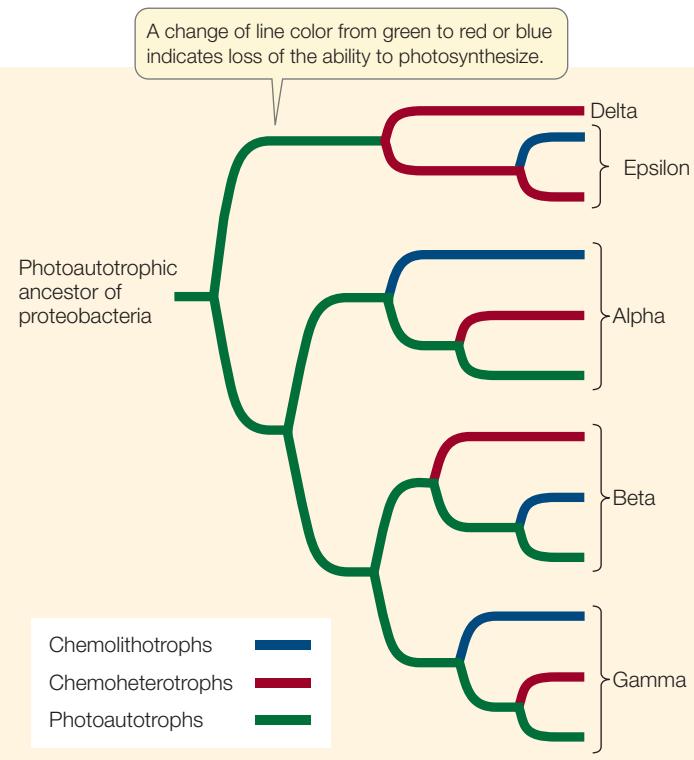
By far the largest group of bacteria, in terms of number of described species, is the **proteobacteria**. The proteobacteria include many species of Gram-negative, bacteriochlorophyll-containing, sulfur-using photoautotrophs—as well as dramatically diverse bacteria that bear no phenotypic resemblance to the photoautotrophic species. Genetic and morphological evidence indicates that the mitochondria of eukaryotes were derived from a proteobacterium by endosymbiosis (see Section 27.2).

No characteristic demonstrates the diversity of the proteobacteria more clearly than their metabolic pathways (Figure 26.19). There are five groups of proteobacteria: alpha, beta, gamma, delta, and epsilon. The common ancestor of all the pro-



26.18 The Tiniest Cells Containing only about one-fifth as much DNA as *E. coli*, mycoplasmas are the smallest known bacteria.

teobacteria was a photoautotroph. Early in evolution, two groups of proteobacteria lost their ability to photosynthesize and have been chemotrophs ever since. The other three groups still have photoautotrophic members, but in each group some evolutionary lines have abandoned photoautotrophy and taken up other modes of nutrition. There are chemolithotrophs and chemoheterotrophs in all three groups. Why? One possibility is that each of the trends shown in Figure 26.19 was an evolution-



26.19 Modes of Nutrition in the Proteobacteria The common ancestor of all proteobacteria was probably a photoautotroph. As they encountered new environments, the delta and epsilon proteobacteria lost the ability to photosynthesize. In the other three groups, some evolutionary lineages became chemolithotrophs or chemoheterotrophs.

ary response to selective pressures encountered as these bacteria colonized new habitats that presented new challenges and opportunities. Lateral gene transfer may have played a role in these responses.

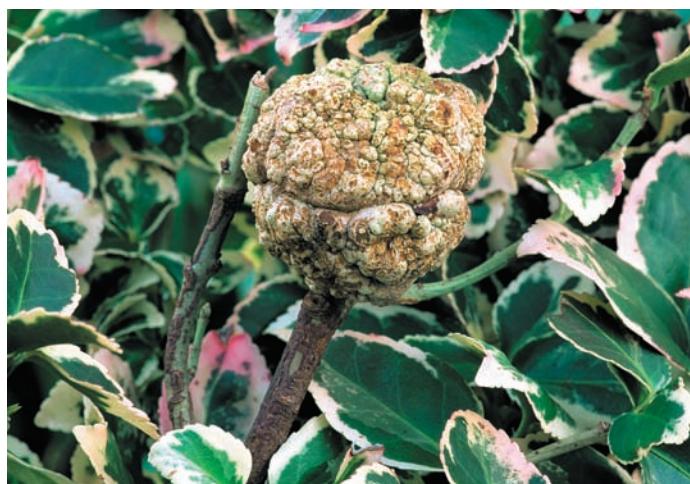
Among the proteobacteria are some nitrogen-fixing genera, such as *Rhizobium* (see Figure 36.9), and other bacteria that contribute to the global nitrogen and sulfur cycles. *Escherichia coli*, one of the most studied organisms on Earth, is a proteobacterium. So, too, are many of the most famous human pathogens, such as *Yersinia pestis* (which causes bubonic plague), *Vibrio cholerae* (cholera), and *Salmonella typhimurium* (gastrointestinal disease).

Although fungi cause most plant diseases, and viruses cause others, about 200 known plant diseases are of bacterial origin. *Crown gall*, with its characteristic tumors (Figure 26.20), is one of the most striking. The causal agent of crown gall is *Agrobacterium tumefaciens*, a proteobacterium that harbors a plasmid used in recombinant DNA studies as a vehicle for inserting genes into new plant hosts (see Section 18.2).

We have discussed six clades of bacteria in some detail. Other bacterial clades are well known, and there are probably dozens more waiting to be discovered. This estimate is conservative because so few bacteria have been cultured and studied in the laboratory.

Archaea differ in several important ways from bacteria

The separation of Archaea from Bacteria and Eukarya was originally based on phylogenetic relationships determined from sequences of rRNA genes. This conclusion was supported when biologists sequenced the first archaeal genome. It consisted of 1,738 genes, more than half of which were unlike any genes ever found in the other two domains. Archaea are well known for living in extreme habitats such as those with high salinity (salt content), low oxygen concentrations, high temperatures, or high or low pH (Figure 26.21). However, many archaea are not extremeophiles but live in moderate habitats; they are common in



26.20 Crown Gall A crown gall, the type of tumor shown here growing on the stem of a bushy shrub, is caused by the proteobacterium *Agrobacterium tumefaciens*.

INVESTIGATING LIFE

26.21 What Is the Highest Temperature Compatible with Life?

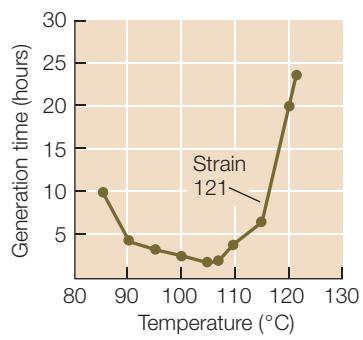
Can any organism thrive at temperatures above 120°C? This is the temperature used for sterilization, known to destroy all previously described organisms. Kazem Kashefi and Derek Lovley isolated an unidentified prokaryote from water samples taken near a hydrothermal vent and found it survived and even grew at 121°C. The organism was dubbed “Strain 121,” and its gene sequencing results indicate that it is an archaeal species.

HYPOTHESIS Some prokaryotes survive and even multiply at temperatures above the 120°C threshold of sterilization.

METHOD

1. Seal samples of unidentified, iron-reducing, thermal vent prokaryotes in tubes with a medium containing Fe³⁺ as an electron acceptor. Control tubes contain Fe³⁺ but no organisms.
2. Hold both tubes in a sterilizer at 121°C for 10 hours. If the iron-reducing organisms are metabolically active, they will reduce the Fe³⁺ to Fe²⁺ (as magnetite, which can be detected with a magnet).
3. Isolate any surviving organisms and test for growth at various temperatures.

RESULTS



The iron-containing solids were attracted to a magnet only in those tubes that contained living cells.

Cells multiplied most rapidly at about 105°C but divided about once a day even at 121°C.

CONCLUSION Some prokaryotic organisms can survive and grow at temperatures above the previously defined sterilization limit.

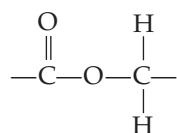
FURTHER INVESTIGATION: Note that Strain 121 did not grow during a 2-hour exposure to 130°C, but it did not die, either. How would you demonstrate that it was still alive?

Go to yourBioPortal.com for original citations, discussions, and relevant links for all INVESTIGATING LIFE figures.

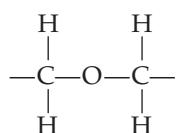
soil, for example. Perhaps the largest number of archaea live in the ocean depths.

One current classification scheme divides Archaea into two principal groups, **Euryarchaeota** and **Crenarchaeota**. Less is known about two more recently discovered groups, **Korarchaeota** and **Nanoarchaeota**. In fact, we know relatively little about the phylogeny of archaea, in part because the study of archaea is still in its early stages.

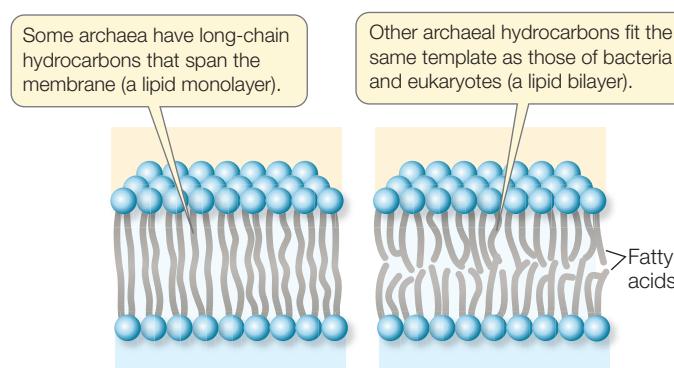
Two characteristics shared by all archaea are the absence of peptidoglycan in their cell walls and the presence of lipids of distinctive composition in their cell membranes (see Table 26.1). The unusual lipids in the membranes of archaea are found in all archaea and in no bacteria or eukaryotes. Most bacterial and eukaryotic membrane lipids contain unbranched long-chain fatty acids connected to glycerol molecules by *ester linkages*:



In contrast, some archaeal membrane lipids contain long-chain hydrocarbons connected to glycerol molecules by *ether linkages*:



These ether linkages are a synapomorphy of archaea. In addition, the long-chain hydrocarbons of archaea are branched. One class of these lipids, with hydrocarbon chains 40 carbon atoms in length, contains glycerol at *both* ends of the hydrocarbons (**Figure 26.22**). This *lipid monolayer* structure, unique to archaea, still fits in a biological membrane because the lipids are twice as long as the typical lipids in the bilayers of other membranes. Lipid monolayers and bilayers are both found among the archaea. The effects, if any, of these structural features on membrane performance are unknown. In spite of this striking difference in their membrane lipids, the membranes seen in all three domains have similar overall structures, dimensions, and functions.



26.22 Membrane Architecture in Archaea The long-chain hydrocarbons of many archaeal membranes have glycerol molecules at both ends, so that the membranes consist of a lipid monolayer. In contrast, the membranes of other archaea, bacteria, and eukaryotes consist of a lipid bilayer.

Most Crenarchaeota live in hot and/or acidic places

Most known Crenarchaeota are either thermophilic (heat loving), acidophilic (acid loving), or both. Members of the genus *Sulfolobus* live in hot sulfur springs at temperatures of 70°C to 75°C. They become metabolically inactive at 55°C (131°F). Hot sulfur springs are also extremely acidic. *Sulfolobus* grows best in the range from pH 2 to pH 3, but some members of this genus readily tolerate pH values as low as 0.9. Most acidophilic thermophiles maintain an internal pH of 5.5 to 7 (close to neutral) in spite of their acidic environment. These and other thermophiles thrive where very few other organisms can even survive. The archaea living in the volcanic vent shown in the opening of this chapter are examples of such thermophiles.

Euryarchaeota are found in surprising places

Some species of Euryarchaeota share the property of producing methane (CH_4) by reducing carbon dioxide. All of these methanogens are obligate anaerobes, and methane production is the key step in their energy metabolism. Comparison of rRNA gene sequences has revealed a close evolutionary relationship among these methanogenic species, which were previously assigned to several different bacterial groups.

Methanogens release approximately 2 billion tons of methane gas into Earth's atmosphere each year, accounting for 80 to 90 percent of the methane in the atmosphere, including the methane produced in some mammalian digestive systems. Approximately a third of this methane comes from methanogens living in the guts of grazing herbivores such as cattle, sheep, and deer, and another large fraction comes from the methanogens that live in the guts of termites and cockroaches. Methane is increasing in Earth's atmosphere by about 1 percent per year and contributes to the greenhouse effect. Part of the increase is due to increases in cattle and rice farming and the methanogens associated with both.

One methanogen, *Methanopyrus*, lives on the ocean bottom near hot hydrothermal vents. *Methanopyrus* can survive and grow at 122°C. It grows best at 98°C and not at all at temperatures below 84°C.

Another group of Euryarchaeota, the **extreme halophiles** (salt lovers), lives exclusively in very salty environments, such as the water of Blood Falls described in the opening of this chapter. Because they contain pink carotenoid pigments, halophiles are sometimes easy to see (**Figure 26.23**). Halophiles grow in the Dead Sea and in brines of all types: the reddish pink spots that can occur on pickled fish are colonies of halophilic archaea. Few other organisms can live in the saltiest of the homes that the extreme halophiles occupy; most would "dry" to death, losing too much water to the hypertonic environment. Extreme halophiles have been found in lakes with pH values as high as 11.5—the most alkaline environment inhabited by living organisms, and almost as alkaline as household ammonia.

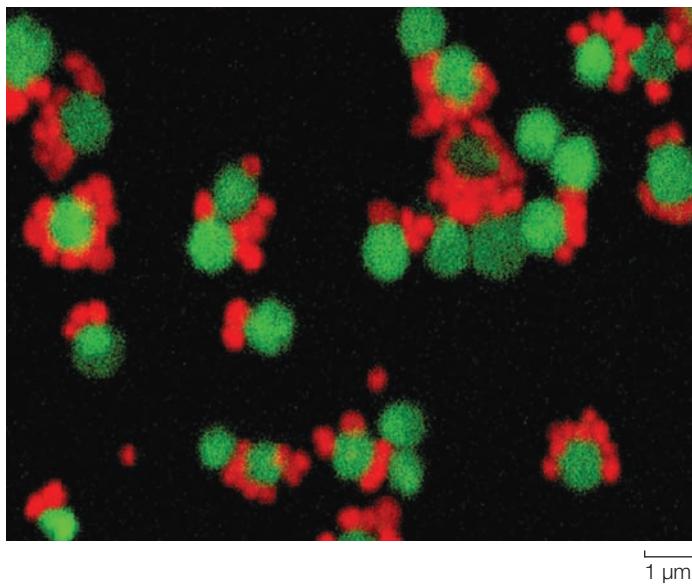
Some of the extreme halophiles have a unique system for trapping light energy and using it to form ATP—without using any form of chlorophyll—when oxygen is in short supply. They use the pigment *retinal* (also found in the vertebrate eye) com-



26.23 Extreme Halophiles Commercial seawater evaporating ponds (these are in San Francisco Bay) are home to salt-loving archaea, easily visible here because of their carotenoid pigments.

bined with a protein to form a light-absorbing molecule called *bacteriorhodopsin*, and they form ATP by a chemiosmotic mechanism of the kind described in Figure 9.9.

Another member of the Euryarchaeota, *Thermoplasma*, has no cell wall. It is thermophilic and acidophilic, its metabolism is aerobic, and it lives in coal deposits. It has the smallest genome among the archaea, and among the smallest (along with the mycoplasmas) of any free-living organism—1,100,000 base pairs.



26.24 A Nanoarchaeote Growing in Mixed Culture with a Crenarchaeote *Nanoarchaeum equitans* (red), discovered living near deep-ocean hydrothermal vents, is the only representative of the nanoarchaeote group so far discovered. This tiny organism lives attached to cells of the crenarchaeote *Ignicoccus* (green). For this confocal laser micrograph, the two species were visually differentiated by fluorescent dye “tags” that are specific to their separate gene sequences.

Korarchaeota and Nanoarchaeota are less well known

The Korarchaeota are known only by evidence derived from DNA isolated directly from hot springs. No korarchaeote has been successfully grown in pure culture.

Another distinctive archaeal lineage has been discovered at a deep-sea hydrothermal vent off the coast of Iceland. It is the first representative of a group christened Nanoarchaeota because of their minute size. This organism lives attached to cells of *Ignicoccus*, a crenarchaeote. Because of their association, the two species can be grown together in culture (Figure 26.24).

26.4 RECAP

Bacteria and Archaea are highly diverse groups that survive in almost every imaginable habitat on Earth. Many can survive and even thrive in habitats where no eukaryotes can live, including extremely hot, acidic, or alkaline conditions.

- Can you explain how metabolic diversity could have become so great in the proteobacteria? See pp. 550–551 and Figure 26.19
- What makes the membranes of archaea unique? See p. 552 and Figure 26.22

Because prokaryotes have so many different metabolic and nutritional capabilities, and because they can live in so many environments, it is reasonable to expect that they affect their environments in many ways. As we are about to see, prokaryotes directly affect humans—in ways both beneficial and harmful.

26.5 How Do Prokaryotes Affect Their Environments?

Prokaryotes live in and exploit all kinds of environments and are part of all ecosystems. In this section we examine the roles of prokaryotes that live in soils, in water, and even in other organisms, where they may exist in a neutral, beneficial, or parasitic relationship with their host’s tissues. The roles of some prokaryotes living in extreme environments have yet to be determined.

Remember that in spite of our frequent mention of prokaryotes as human pathogens, only a small minority of the known prokaryotic species are pathogenic. Many more prokaryotes play positive roles in our lives and in the biosphere. We make direct use of many bacteria and a few archaea in such diverse applications as cheese production, sewage treatment, and the industrial production of an amazing variety of antibiotics, vitamins, organic solvents, and other chemicals.

Prokaryotes are important players in element cycling

Many prokaryotes are decomposers—organisms that metabolize organic compounds in dead organisms and other organic material and return the products to the environment as inorganic substances. Prokaryotes, along with fungi, return tremen-

dous quantities of organic carbon to the atmosphere as carbon dioxide, thus carrying out a key step in the carbon cycle. Prokaryotic decomposers also return inorganic nitrogen and sulfur to the environment.

Animals depend on plants and other photosynthetic organisms for their food, directly or indirectly. But plants depend on other organisms—prokaryotes—for their own nutrition. The extent and diversity of life on Earth would not be possible without nitrogen fixation by prokaryotes. Nitrifiers are crucial to the biosphere because they convert the products of nitrogen fixation into nitrate ions, the form of nitrogen most easily used by many plants (see Figure 36.11). Plants, in turn, are the source of nitrogen compounds for animals and fungi. Denitrifiers also play a key role in keeping the nitrogen cycle going. Without denitrifiers, which convert nitrate ions back into nitrogen gas, all forms of nitrogen would leach from the soil and end up in lakes and oceans, making life on land impossible. Other prokaryotes—both bacteria and archaea—contribute to a similar cycle of sulfur.

In the ancient past, the cyanobacteria had an equally dramatic effect on life: their photosynthesis generated oxygen, converting Earth's atmosphere from an anaerobic to an aerobic environment (see Section 25.3). A major result was the wholesale loss of obligate anaerobic species that could not tolerate the oxygen generated by the cyanobacteria. Only those anaerobes that adapted to aerobic conditions or colonized environments that remained anaerobic survived. However, this transformation to aerobic environments made possible the evolution of cellular respiration and the subsequent explosion of eukaryotic life.

Prokaryotes live on and in other organisms

Prokaryotes work together with eukaryotes in many ways. As we have seen, mitochondria and chloroplasts are descended from what were once free-living bacteria. Much later in evolutionary history, some plants became associated with bacteria to form cooperative nitrogen-fixing nodules on their roots (see Figure 36.9).

Many animals harbor a variety of bacteria and archaea in their digestive tracts. Cattle depend on prokaryotes to perform important steps in digestion. Like most animals, cattle cannot produce cellulase, the enzyme needed to start the digestion of the cellulose that makes up the bulk of their plant food. However, bacteria living in a special section of the gut, called the rumen, produce enough cellulase to process the daily diet for the cattle.

Humans use some of the metabolic products—especially vitamins B₁₂ and K—of bacteria living in our large intestine. These and other bacteria and archaea line our intestines with a dense biofilm that is in intimate contact with the mucosal lining of the gut. This biofilm facilitates nutrient transfer from the intestine into the body and induces immunity to the gut contents. The biofilm in the gut is a major part of an “organ” consisting of prokaryotes that is essential to our health. Its makeup varies from time to time and from region to region of the intestinal tract, and it has a complex ecology that scientists have just begun to explore in detail—including the possibility that the species composition of an individual's prokaryote gut fauna may contribute to obesity (or the resistance to it).

We are heavily populated inside and out by bacteria. A 2009 study of bacteria that live on human skin identified more than 1,000 species living on the outside of our bodies, and many of these are thought to be critical to maintaining the skin's health. Although only a small percentage of bacterial species are agents of disease, popular notions of bacteria as “germs” and fear of the consequences of infection arouse our curiosity about those few.

A small minority of bacteria are pathogens

The late nineteenth century was a productive era in the history of medicine—a time when bacteriologists, chemists, and physicians proved that many diseases are caused by microbial agents. During this time the German physician Robert Koch laid down a set of four rules for establishing that a particular microorganism causes a particular disease:

- The microorganism is always found in individuals with the disease.
- The microorganism can be taken from the host and grown in pure culture.
- A sample of the culture produces the same disease when injected into a new, healthy host.
- The newly infected host yields a new, pure culture of microorganisms identical to those obtained in the second step.

These rules, called **Koch's postulates**, were very important in a time when it was not widely understood that microorganisms cause disease. Although medical science today has more powerful diagnostic tools, the postulates remain useful on occasion. For example, physicians were taken aback in the 1990s when stomach ulcers—long accepted and treated as the result of excess stomach acid—were shown by Koch's postulates to be caused by the bacterium *Helicobacter pylori* (see Figure 51.14).

Only a tiny percentage of all prokaryotes are pathogens, and of those that are known, all are in the domain Bacteria. For an organism to be a successful pathogen, it must:

- arrive at the body surface of a potential host;
- enter the host's body;
- evade the host's defenses;
- multiply inside the host; and finally
- infect a new host.

Failure to successfully complete any of these steps ends the reproductive career of a pathogenic organism. However, in spite of the many defenses available to potential hosts (see Chapter 42), some bacteria are very successful pathogens.

For the host, the consequences of a bacterial infection depend on several factors. One is the **invasiveness** of the pathogen—its ability to multiply in the host's body. Another is its **toxigenicity**—its ability to produce toxins, chemical substances that are harmful to the host's tissues. *Corynebacterium diphtheriae*, the agent that causes diphtheria, has low invasiveness and multiplies only in the throat, but its toxigenicity is so great that the entire body is affected. In contrast, *Bacillus anthracis*, which causes anthrax (a disease primarily of cattle and sheep, but

which is also sometimes fatal in humans), has low toxigenicity but is so invasive that the entire bloodstream ultimately teems with the bacteria.

There are two general types of bacterial toxins: exotoxins and endotoxins. **Endotoxins** are released when certain Gram-negative bacteria grow or lyse (burst). Endotoxins are lipopolysaccharides (complexes consisting of a polysaccharide and a lipid component) that form part of the outer bacterial membrane. Endotoxins are rarely fatal; they normally cause fever, vomiting, and diarrhea. Among the endotoxin producers are some strains of the gamma-proteobacteria *Salmonella* and *Escherichia*.

Exotoxins are usually soluble proteins released by living, multiplying bacteria, and they may travel throughout the host's body. They are highly toxic—often fatal—to the host, but they do not produce fevers. Human diseases induced by bacterial exotoxins include tetanus (*Clostridium tetani*), cholera (*Vibrio cholerae*), and bubonic plague (*Yersinia pestis*). Anthrax results from three exotoxins produced by *Bacillus anthracis*. Botulism is caused by exotoxins produced by *Clostridium botulinum* that are among the most poisonous ever discovered. The lethal dose of the botulinum A exotoxin for humans is about one-millionth of a gram (1 µg). Nonetheless, much smaller doses of this exotoxin, marketed under various trade names, are used to treat muscle spasms and also for cosmetic purposes (temporary wrinkle reduction in skin).

Pathogenic bacteria are often surprisingly difficult to combat, even with today's arsenal of antibiotics. One source of this difficulty is the ability of prokaryotes to form biofilms.

26.5 RECAP

Prokaryotes play key roles in the cycling of Earth's elements. Many prokaryotes are beneficial and even necessary to other forms of life; others are pathogens.

- Describe the roles of bacteria in the nitrogen cycle.
See p. 554
- What are some of the challenges facing a pathogen?
See p. 554

Before moving on to discuss the diversity of eukaryotic life, it is appropriate to consider how viruses are related to the rest of life. Although they are not cellular, viruses are numerically among the most abundant organisms on Earth. Their effects on other organisms are enormous. Where did viruses come from, and how do they fit into the tree of life? Biologists are still working to answer these questions.

26.6 Where Do Viruses Fit into the Tree of Life?

Some biologists do not think of viruses as living organisms, primarily because they are not cellular and must depend on cellular organisms for basic life functions such as replication and metabolism. But viruses are derived from the cells of other living

organisms. They use the same essential forms of genetic storage and transmission as do cellular organisms. Viruses infect all cellular forms of life, including bacteria, archaea, and eukaryotes. They replicate, mutate, evolve, and interact with other organisms, often causing serious diseases when they infect their hosts. They are also numerically among the most abundant organisms on the planet. And, finally, viruses clearly evolve independently of other organisms, so it is almost impossible not to treat them as a part of life.

Several factors make virus phylogeny difficult to resolve. The tiny size of many viral genomes restricts the phylogenetic analyses that can be conducted to relate viruses to cellular organisms. The rapid mutational rate, which results in rapid evolution of viral genomes, tends to cloud evolutionary relationships across long periods of time. There are no known viral fossils (viruses are too small and delicate to fossilize), so the paleontological record offers no clues as to viral origins. Finally, viruses are highly diverse (see Figure 26.25), and several lines of evidence support the hypothesis that viruses have evolved repeatedly within each of the major groups of life.

Many RNA viruses probably represent escaped genomic components

Although viruses are now obligate parasites of cellular species, they may once have been cellular components involved in basic cellular functions—that is, they may be “escaped” components of cellular life that now evolve independently of their hosts.

NEGATIVE-SENSE SINGLE-STRANDED RNA VIRUSES A case in point is a class of viruses whose genome is composed of single-stranded RNA that is the complement (negative-sense) of the mRNA needed for protein translation. Many of these negative-sense single-stranded RNA viruses have only a few genes, including an RNA-dependent RNA polymerase that allows them to make mRNA from their negative-sense RNA genome. Modern cellular organisms cannot generate mRNA in this manner (at least in the absence of viral infections), but scientists speculate that single-stranded RNA genomes may have been common in the distant past, before DNA became the primary molecule for genetic information storage.

A self-replicating RNA polymerase gene that begins to replicate independently of a cellular genome could conceivably acquire a few additional protein-coding genes through recombination with its host's DNA. If one or more of these genes were to foster the development of a protein coat, the virus might then survive outside the host and infect new hosts. It is believed that this scenario has been repeated many times independently across the tree of life, given that many of the negative-sense single-stranded RNA viruses that infect organisms from bacteria to humans are not closely related to one another. In other words, negative-sense single-stranded RNA viruses do not represent a distinct taxonomic group, but exemplify a particular process of cellular escape that probably happened many different times.

Examples of familiar negative-sense single-stranded RNA viruses include the viruses that cause measles, mumps, rabies, and influenza (Figure 26.25A).