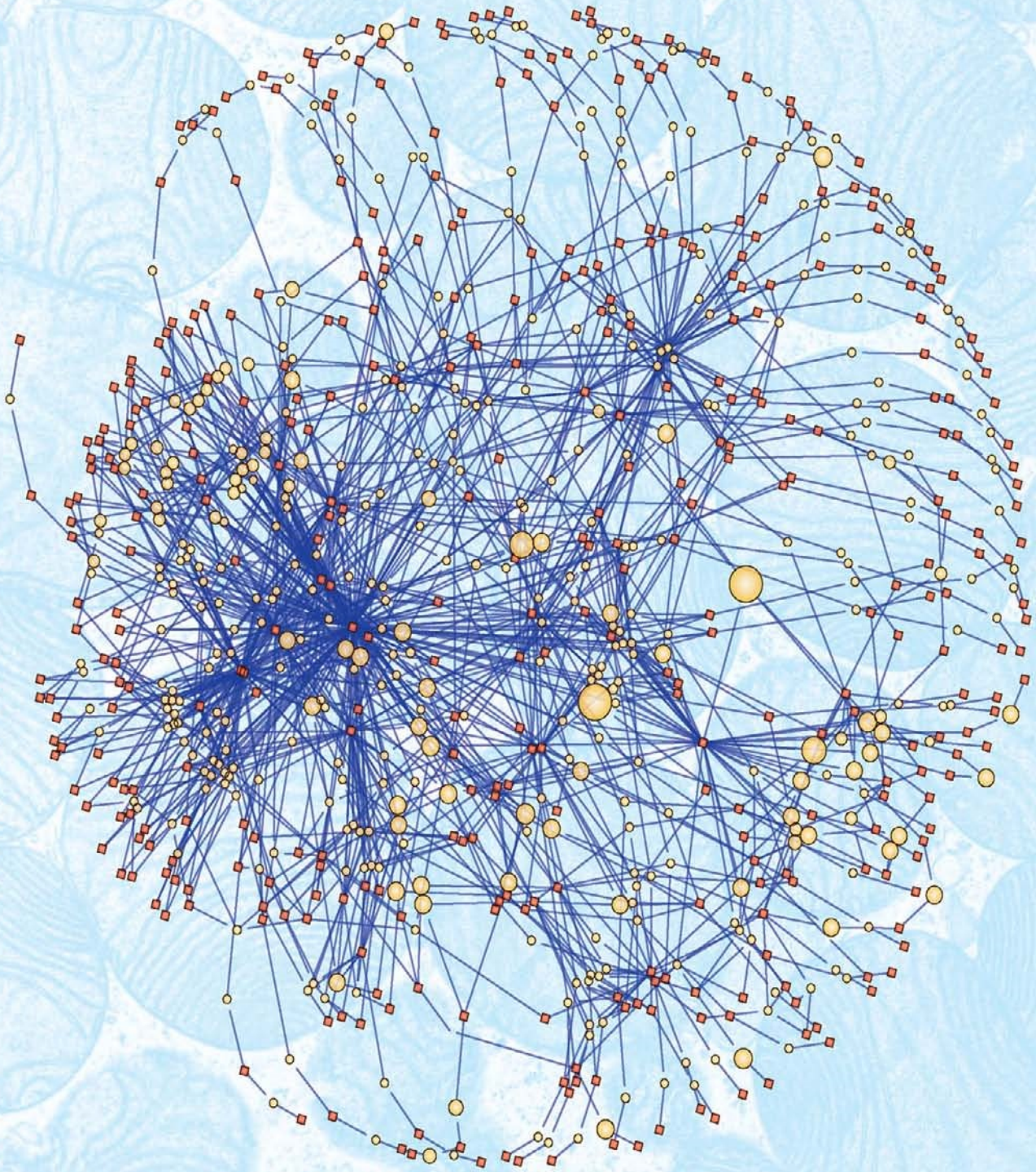
Lehninger

SIXTH EDITION

Principles of Biochemistry

David L. Nelson | Michael M. Cox



Media Connections

Below is a chapter-by-chapter list of the media resources available on the Instructor's CD-ROM and website www.courses.bfwpub.com/lehninger6e.

- Mechanism Animations (12 total) show key reactions in detail.
- Technique Animations (10 total) reveal the experimental techniques available to researchers today.
- Living Graphs (15 total) allow students to alter the parameters in key equations and graph the results.
- Molecular Structure Tutorials (9 total) guide students through concepts using three-dimensional molecular models.

New animations will be added throughout the life of the edition.

Chapter 2 Water

Living Graph: Henderson-Hasselbalch Equation

Chapter 3 Amino Acids, Peptides, and Proteins

Molecular Structure Tutorials: Protein Architecture—Amino Acids

Technique Animation: SDS Gel Electrophoresis

Chapter 4 The Three-Dimensional Structure of Proteins

Molecular Structure Tutorials:
Protein Architecture—Sequence and Primary Structure

Protein Architecture—The α Helix

Protein Architecture—The β Sheet

Protein Architecture—Turn

Protein Architecture—Introduction to Tertiary Structure

Protein Architecture—Tertiary Structure of Fibrous Proteins

Protein Architecture—Tertiary Structure of Small Globular Proteins

Protein Architecture—Tertiary Structure of Large Globular Proteins

Protein Architecture—Quaternary Structure

Chapter 5 Protein Function

Molecular Structure Tutorial: Oxygen-Binding Proteins—Myoglobin: Oxygen Storage

Living Graphs:
Protein-Ligand Interactions

Binding Curve for Myoglobin

Molecular Structure Tutorial: Oxygen-Binding Proteins—Hemoglobin: Oxygen Transport

Living Graphs:
Cooperative Ligand Binding

Hill Equation

Molecular Structure Tutorials:
Oxygen-Binding Proteins—Hemoglobin Is Susceptible to Allosteric Regulation

Oxygen-Binding Proteins—Defects in Hb Lead to Serious Genetic Disease

MHC Molecules

Technique Animation: Immunoblotting

Chapter 6 Enzymes

Living Graphs:
Michaelis-Menten Equation

Competitive Inhibitor

Uncompetitive Inhibitor

Mixed Inhibitor

Mechanism Animation: Chymotrypsin Mechanism

Living Graph: Lineweaver-Burk Equation

Chapter 8 Nucleotides and Nucleic Acids

Molecular Structure Tutorial: Nucleotides, Building Blocks of Amino Acids

Technique Animation: Dideoxy Sequencing of DNA

Chapter 9 DNA-Based Information Technologies

Molecular Structure Tutorial: Restriction Endonucleases

Technique Animations:
Plasmid Cloning

Reporter Constructs

Polymerase Chain Reaction

Synthesizing an Oligonucleotide Array

Screening an Oligonucleotide Array for Patterns of Gene Expression

Yeast Two-Hybrid Systems

Creating a Transgenic Mouse

Chapter 11 Biological Membranes and Transport

Living Graphs:

Free-Energy Change for Transport

Free-Energy Change for Transport of an Ion

Chapter 12 Biosignaling

Molecular Structure Tutorial: Trimeric G Proteins—Molecular On/Off Switches

Chapter 13 Bioenergetics and Biochemical Reaction Types

Living Graphs:

Free-Energy Change

Free-Energy of Hydrolysis of ATP

Chapter 14 Glycolysis, Gluconeogenesis, and the Pentose Phosphate Pathway

Mechanism Animations:

Phosphohexose Isomerase Mechanism

Alcohol Dehydrogenase Mechanism

Thiamine Pyrophosphate Mechanism

Chapter 16 The Citric Acid Cycle

Mechanism Animation: Citrate Synthase Mechanism

Chapter 17 Fatty Acid Catabolism

Mechanism Animation: Fatty Acyl–CoA Synthetase Mechanism

Chapter 18 Amino Acid Oxidation and the Production of Urea

Mechanism Animations:

Pyridoxal Phosphate Reaction Mechanism

Carbamoyl Phosphate Synthetase I Mechanism

Argininosuccinate Synthetase Mechanism

Chapter 19 Oxidative Phosphorylation and Photophosphorylation

Living Graph: Free-Energy Change for Transport of an Ion

Molecular Structure Tutorial: Bacteriorhodopsin

Chapter 20 Carbohydrate Biosynthesis in Plants and Bacteria

Mechanism Animation: Rubisco Mechanism

Chapter 22 Biosynthesis of Amino Acids, Nucleotides, and Related Molecules

Mechanism Animations:

Tryptophan Synthase Mechanism

Thymidylate Synthase Mechanism

Chapter 24 Genes and Chromosomes

Animation: Three-Dimensional Packaging of Nuclear Chromosomes

Chapter 25 DNA Metabolism

Molecular Structure Tutorial: Restriction Endonucleases

Animation:

Nucleotide Polymerization by DNA Polymerase

DNA Synthesis

Chapter 26 RNA Metabolism

Animation: mRNA Splicing

Molecular Structure Tutorial: Hammerhead Ribozyme

Animation: Life Cycle of an mRNA

Chapter 28 Regulation of Gene Expression

Molecular Structure Tutorial: Lac Repressor

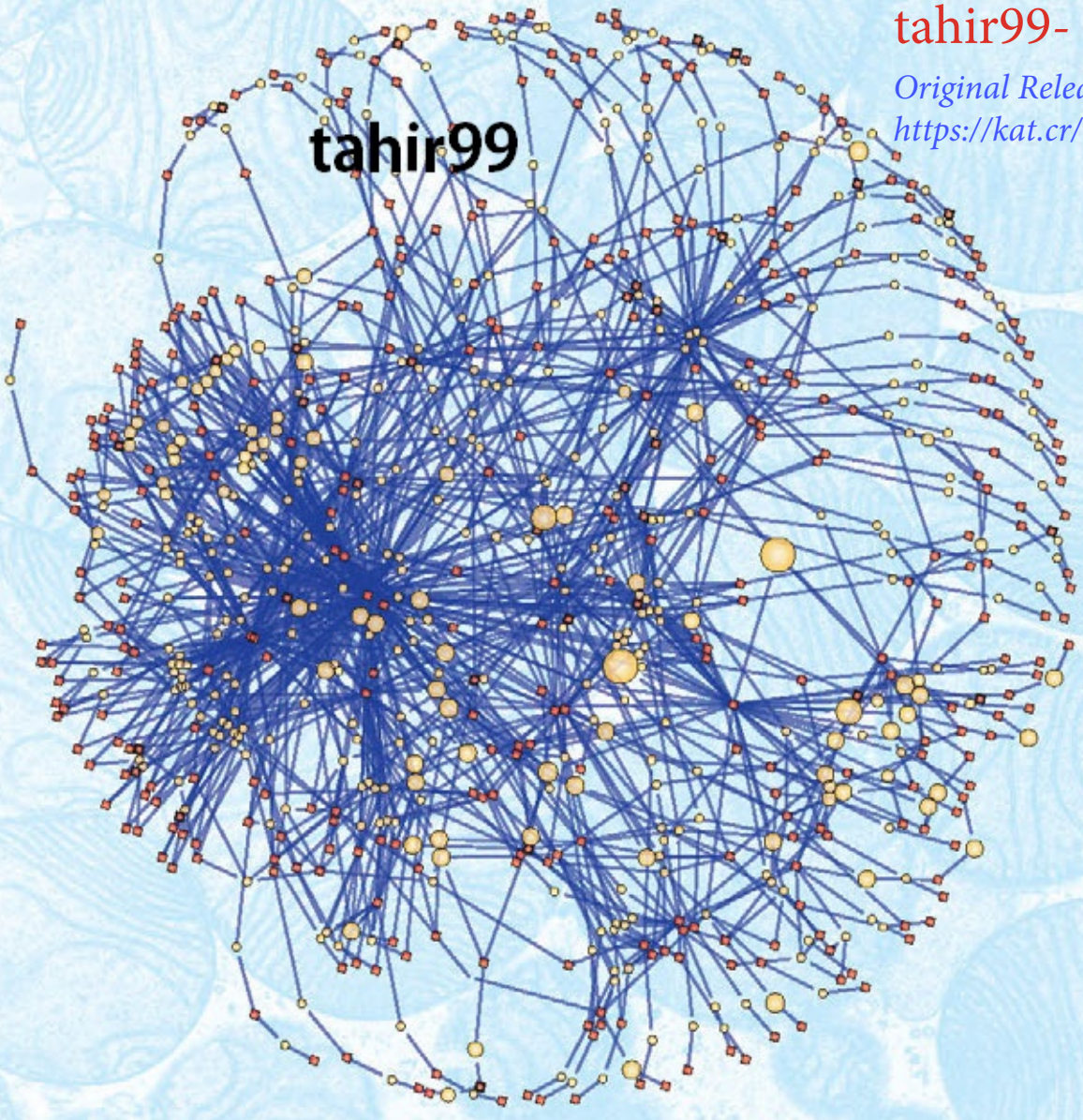
Lehninger

Principles of Biochemistry

SIXTH EDITION

David L. Nelson
Professor of Biochemistry
University of Wisconsin–Madison

Michael M. Cox
Professor of Biochemistry
University of Wisconsin–Madison



W. H. FREEMAN AND COMPANY • New York

Publisher:	SUSAN WINSLOW
Senior Acquisitions Editor:	LAUREN SCHULTZ
Senior Developmental Editor:	SUSAN MORAN
Developmental Editor:	MATTHEW TONTONOZ
Associate Director of Marketing:	DEBBIE CLARE
Marketing Director:	JOHN BRITCH
Marketing Assistant	LINDSAY NEFF
Media Editor:	ALLISON MICHAEL
Managing Editor:	PHILIP McCAFFREY
Project Editor:	JANE O'NEILL
Photo Editor:	TED SZCZEPANSKI
Photo Researcher:	ELYSE RIEDER
Art Director:	DIANA BLUME
Illustration Coordinator:	JANICE DONNOLA
Illustrations:	H. ADAM STEINBERG and DRAGONFLY MEDIA GROUP
Molecular Graphics:	H. ADAM STEINBERG
Production Manager:	SUSAN WEIN
Composition:	APTARA, INC.
Printing and binding:	QUAD/GRAPHICS VERSAILLES

North American Edition
Cover image: The network of interactions in an animal mitochondrion. Each dot represents a compound, and each line, an enzyme that interconverts the two compounds. The major nodes include ADP, ATP, NAD^+, and NADH. The image was constructed with Cytoscape software by Anthony Smith in the laboratory of Alan Robinson, Medical Research Council Mitochondrial Biology Unit, Cambridge, UK, using data from MitoMiner (Smith, A.C., Blackshaw, J.A., & Robinson, A.J. (2012) MitoMiner: a data warehouse for mitochondrial proteomics data. *Nucleic Acids Res*. 40, D1160–D1167). *Background*: Transmission electron micrograph of interscapular brown adipose cell from a bat. (Don W. Fawcett/Science Source/Photo Researchers)

International Edition
Cover design: Dirk Kaufman
Cover image: Nastco/iStockphoto.com

Library of Congress Control Number: 2012948755

North American Edition
ISBN-13: 978-1-4292-3414-6
ISBN-10: 1-4292-3414-8

International Edition
ISBN-13: 978-1-4641-0962-1
ISBN-10: 1-4641-0962-1

Printed in the United States of America

First printing

W. H. Freeman and Company
41 Madison Avenue
New York, NY 10010
www.whfreeman.com

Macmillan Higher Education
Houndmills, Basingstoke
RG21 6XS, England
www.macmillanhighered.com/international

To Our Teachers

Paul R. Burton
Albert Finholt
William P. Jencks
Eugene P. Kennedy
Homer Knoss
Arthur Kornberg
I. Robert Lehman
Earl K. Nelson
Wesley A. Pearson
David E. Sheppard
Harold B. White

About the Authors

David L. Nelson, born in Fairmont, Minnesota, received his BS in Chemistry and Biology from St. Olaf College in 1964 and earned his PhD in Biochemistry at Stanford Medical School under Arthur Kornberg. He was a postdoctoral fellow at the Harvard Medical School with Eugene P. Kennedy, who was one of Albert Lehninger's first graduate students. Nelson joined the faculty of the University of Wisconsin–Madison in 1971 and became a full professor of biochemistry in 1982. He was for eight years the Director of the Center for Biology Education at the University of Wisconsin–Madison.

Nelson's research has focused on the signal transductions that regulate ciliary motion and exocytosis in the protozoan *Paramecium*. The enzymes of signal transductions, including a variety of protein kinases, are primary targets of study. His research group has used enzyme purification, immunological techniques, electron microscopy, genetics, molecular biology, and electrophysiology to study these processes.

Dave Nelson has a distinguished record as a lecturer and research supervisor. For 40 years he has taught an intensive survey of biochemistry for advanced biochemistry undergraduates in the life sciences. He has also taught a survey of biochemistry for nursing students, and graduate courses on membrane structure and function and on molecular neurobiology. He has sponsored numerous PhD, MS, and undergraduate honors theses and has received awards for his outstanding teaching, including the Dreyfus Teacher–Scholar Award, the Atwood Distinguished Professorship, and the Unterkofler Excellence in Teaching Award from the University of Wisconsin System. In 1991–1992 he was a visiting professor of chemistry and biology at Spelman College. His second love is history, and in his dotage he has begun to teach the history of biochemistry to undergraduates and to collect antique scientific instruments for use in a laboratory course he teaches.

David L. Nelson and Michael M. Cox

Michael M. Cox was born in Wilmington, Delaware. In his first biochemistry course, Lehninger's *Biochemistry* was a major influence in refocusing his fascination with biology and inspiring him to pursue a career in biochemistry. After graduating from the University of Delaware in 1974, Cox went to Brandeis University to do his doctoral work with William P. Jencks, and then to Stanford in 1979 for postdoctoral study with I. Robert Lehman. He moved to the University of Wisconsin–Madison in 1983 and became a full professor of biochemistry in 1992.

Cox's doctoral research was on general acid and base catalysis as a model for enzyme-catalyzed reactions. At Stanford, he began work on the enzymes involved in genetic recombination. The work focused particularly on the RecA protein, designing purification and assay methods that are still in use, and illuminating the process of DNA branch migration. Exploration of the enzymes of genetic recombination has remained the central theme of his research.

Mike Cox has coordinated a large and active research team at Wisconsin, investigating the enzymology, topology, and energetics of genetic recombination. A primary focus has been the mechanism of RecA protein–mediated DNA strand exchange, the role of ATP in the RecA system, and the regulation of recombinational DNA repair. Part of the research program now focuses on organisms that exhibit an especially robust capacity for DNA repair, such as *Deinococcus radiodurans*, and the applications of those repair systems to biotechnology.

For almost 30 years he has taught (with Dave Nelson) the survey of biochemistry to undergraduates and has lectured in graduate courses on DNA structure and topology, protein-DNA interactions, and the biochemistry of recombination. More recent projects have been the organization of a new course on professional responsibility for first-year graduate students and the establishment of a systematic program to draw talented biochemistry undergraduates into the laboratory at an early stage of their collegiate career. He has received awards for both his teaching and his research, including the Dreyfus Teacher–Scholar Award, the 1989 Eli Lilly Award in Biological Chemistry, and the 2009 Regents Teaching Excellence Award from the University of Wisconsin. He is also highly active in national efforts to provide new guidelines for undergraduate biochemistry education. His hobbies include turning 18 acres of Wisconsin farmland into an arboretum, wine collecting, and assisting in the design of laboratory buildings.

A Note on the Nature of Science

In this twenty-first century, a typical science education often leaves the philosophical underpinnings of science unstated, or relies on oversimplified definitions. As you contemplate a career in science, it may be useful to consider once again the terms **science**, **scientist**, and **scientific method**.

Science is both a way of thinking about the natural world and the sum of the information and theory that result from such thinking. The power and success of science flow directly from its reliance on ideas that can be tested: information on natural phenomena that can be observed, measured, and reproduced and theories that have predictive value. The progress of science rests on a foundational assumption that is often unstated but crucial to the enterprise: that the laws governing forces and phenomena existing in the universe are not subject to change. The Nobel laureate Jacques Monod referred to this underlying assumption as the "postulate of objectivity." The natural world can therefore be understood by applying a process of inquiry—the scientific method. Science could not succeed in a universe that played tricks on us. Other than the postulate of objectivity, science makes no inviolate assumptions about the natural world. A useful scientific idea is one that (1) has been or can be reproducibly substantiated and (2) can be used to accurately predict new phenomena.

Scientific ideas take many forms. The terms that scientists use to describe these forms have meanings quite different from those applied by nonscientists. A *hypothesis* is an idea or assumption that provides a reasonable and testable explanation for one or more observations, but it may lack extensive experimental substantiation. A *scientific theory* is much more than a hunch. It is an idea that has been substantiated to some extent and provides an explanation for a body of experimental observations. A theory can be tested and built upon and is thus a basis for further advance and innovation. When a scientific theory has been repeatedly tested and validated on many fronts, it can be accepted as a fact.

In one important sense, what constitutes science or a scientific idea is defined by whether or not it is published in the scientific literature after peer review by other working scientists. About 16,000 peer-reviewed scientific journals worldwide publish some 1.4 million articles each year, a continuing rich harvest of information that is the birthright of every human being.

Scientists are individuals who rigorously apply the scientific method to understand the natural world. Merely having an advanced degree in a scientific discipline does not make one a scientist, nor does the lack of such a degree prevent one from making important scientific contributions. A scientist must be willing to challenge any idea when new findings demand it. The ideas that a scientist accepts must be based on measurable, reproducible observations, and the scientist must report these observations with complete honesty.

The **scientific method** is actually a collection of paths, all of which may lead to scientific discovery. In the *hypothesis and experiment* path, a scientist poses a hypothesis, then subjects it to experimental test. Many of the processes that biochemists work with every day were discovered in this manner. The DNA structure elucidated by James Watson and Francis Crick led to the hypothesis that base pairing is the basis for information transfer in polynucleotide synthesis. This hypothesis helped inspire the discovery of DNA and RNA polymerases.

Watson and Crick produced their DNA structure through a process of *model building and calculation*. No actual experiments were involved, although the model building and calculations used data collected by other scientists. Many adventurous scientists have applied the process of *exploration and observation* as a path to discovery. Historical voyages of discovery (Charles Darwin's 1831 voyage on H.M.S. *Beagle* among them) helped to map the planet, catalog its living occupants, and change the way we view the world. Modern scientists follow a similar path when they explore the ocean depths or launch probes to other planets. An analog of hypothesis and experiment is *hypothesis and deduction*. Crick reasoned that there must be an adaptor molecule that facilitated translation of the information in messenger RNA into protein. This adaptor hypothesis led to the discovery of transfer RNA by Mahlon Hoagland and Paul Zamecnik.

Not all paths to discovery involve planning. *Serendipity* often plays a role. The discovery of penicillin by Alexander Fleming in 1928 and of RNA catalysts by Thomas Cech in the early 1980s were both chance discoveries, albeit by scientists well prepared to exploit them. *Inspiration* can also lead to important advances. The polymerase chain reaction (PCR), now a central part of biotechnology, was developed by Kary Mullis after a flash of inspiration during a road trip in northern California in 1983.

These many paths to scientific discovery can seem quite different, but they have some important things in common. They are focused on the natural world. They rely on *reproducible observation* and/or *experiment*. All of the ideas, insights, and experimental facts that arise from these endeavors can be tested and reproduced by scientists anywhere in the world. All can be used by other scientists to build new hypotheses and make new discoveries. All lead to information that is properly included in the realm of science. Understanding our universe requires hard work. At the same time, no human endeavor is more exciting and potentially rewarding than trying, and occasionally succeeding, to understand some part of the natural world.

Preface

As we complete our work on this sixth edition of *Lehninger Principles of Biochemistry*, we are again struck by the remarkable changes in the field of biochemistry that have occurred between editions. The sheer volume of new information from high-throughput DNA sequencing, x-ray crystallography, and the manipulation of genes and gene expression, to cite only three examples, challenges both the seasoned researcher and the first-time biochemistry student. Our goal here is to strike a balance: to include new and exciting research findings without making the book overwhelming for students. The primary criterion for inclusion is that the new finding helps to illustrate an important *principle of biochemistry*.

The image on our cover, a map of the known metabolic transformations in a mitochondrion, illustrates the richness of factual material now available about biochemical transformations. We can no longer treat metabolic "pathways" as though they occurred in isolation; a single metabolite may be simultaneously part of many pathways in a three-dimensional network of metabolic transformations. Biochemical research focuses more and more upon the interactions among these pathways, the regulation of their interactions at the level of gene and protein, and the effects of regulation upon the activities of a whole cell or organism.

This edition of *LPOB* reflects these realities. Much of the new material that we have added reflects our increasingly sophisticated understanding of regulatory mechanisms, including those involved in altering the synthesis of enzymes and their degradation, those responsible for the control and timing of DNA synthesis and the cell cycle, and those that integrate the metabolism of carbohydrates, fats, and proteins over time in response to changes in the environment and in different cell types.

Even as we strive to incorporate the latest major advances, certain hallmarks of the book remain unchanged. We continue to emphasize the relevance of biochemistry to the molecular mechanisms of disease, highlighting the special role that biochemistry plays in advancing human health and welfare. A special theme is the metabolic basis of diabetes and the factors that predispose to the disease. This theme is interwoven through many chapters and serves to integrate the discussion of metabolism. We also underscore the importance of evolution to biochemistry. Evolutionary theory is the bedrock upon which all biological sciences rest, and we have not wasted opportunities to highlight its important role in our discipline.

To a significant degree, research progress in biochemistry runs in parallel with the development of better tools and techniques. We have therefore highlighted some of these crucial developments. Chapter 9, DNA-Based Information Technologies, in particular, has been significantly revised to include the latest advances in genomics and next-generation sequencing.

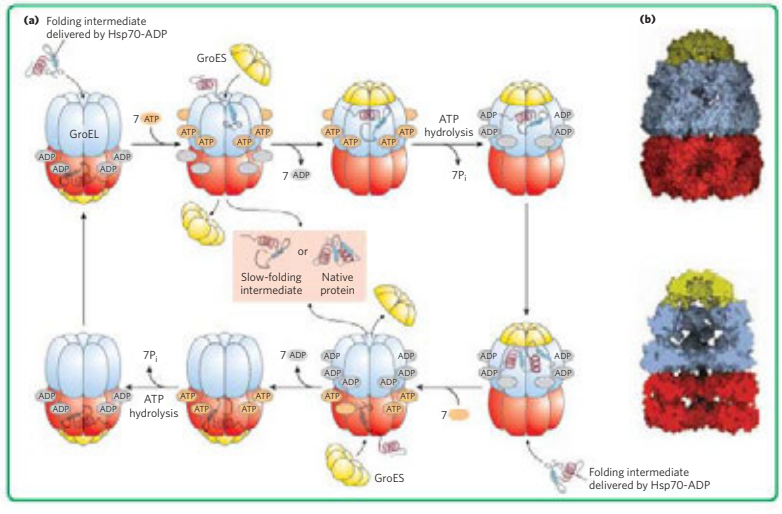
Finally, we have devoted considerable attention to making the text and the art even more useful to students learning biochemistry for the first time. To those familiar with the book, some of these changes will be obvious as soon as you crack the cover.

With every revision of this textbook, we have striven to maintain the qualities that made the original Lehninger text a classic—clear writing, careful explanations of difficult concepts, and insightful communication to students of the ways in which biochemistry is understood and practiced today. The authors have written together for almost 25 years and taught introductory biochemistry together for nearly 30. Our thousands of students at the University of Wisconsin–Madison over those years have been an endless source of ideas about how to present biochemistry more clearly; they have enlightened and inspired us. We hope that this sixth edition of *Lehninger* will in turn enlighten and inspire current students of biochemistry everywhere, and perhaps lead some of them to love biochemistry as we do.

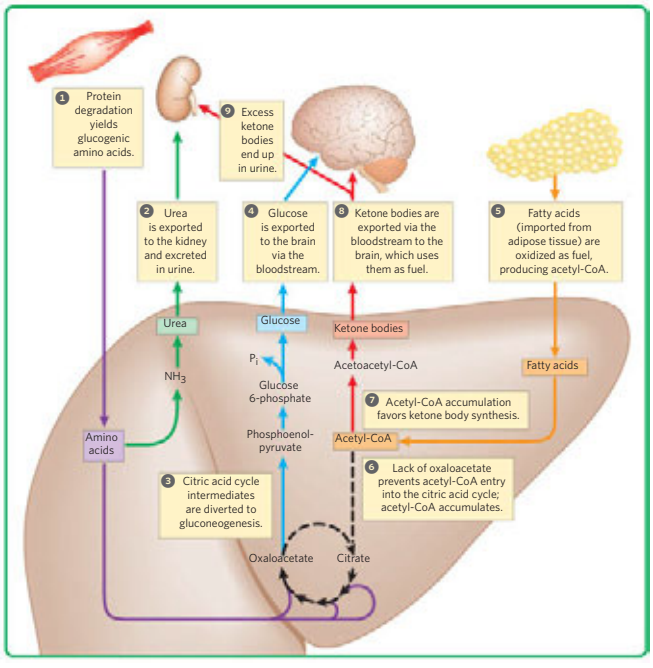
New Art

The most obvious change to the book is the completely revamped art program. Our goal throughout has been to improve pedagogy, drawing on modern graphic resources to make our subject as clear as humanly possible. Many figures illustrate new topics, and much of the art has been reconceived and modernized in style. Defining features of the new art program include:

- **Smarter renditions of classic figures** are easier to interpret and learn from;



Chaperonins in protein folding



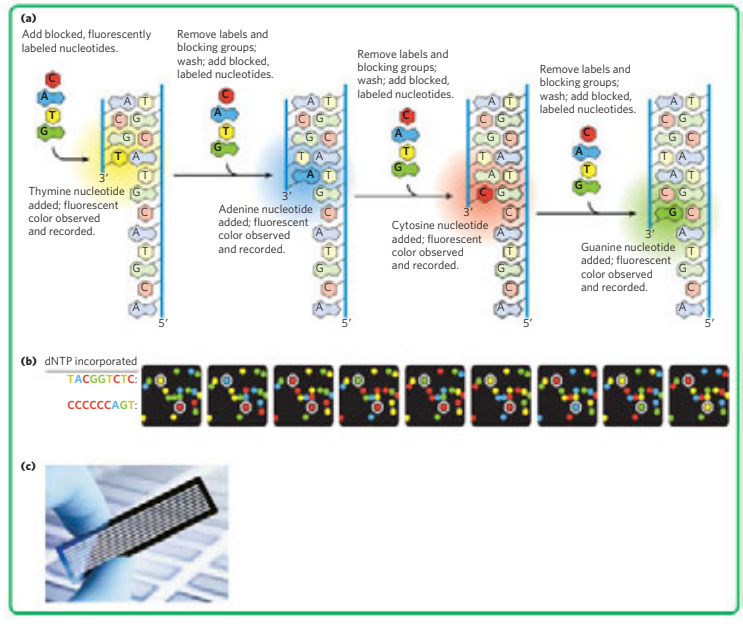
Fuel metabolism in the liver during prolonged fasting or in uncontrolled diabetes mellitus

- **Figures that pair molecular models with schematic cartoons**, generated specifically for this book, use shapes and color schemes that are internally consistent;
- Figures with **numbered, annotated steps** help explain complex processes; in many cases, we have moved descriptive text out of the legends and into the figure itself;
- **Summary figures** help the student to keep the big picture in mind while learning the specifics.

Updated Genomics

Modern genomic techniques have transformed our understanding of biochemistry. In this edition, we have dramatically updated our coverage of genomic methods and their applications. Chapter 9, DNA-Based Information Technologies, has been completely revised to incorporate the latest genomic methods. Many other chapters have been updated to reflect advances gained from these methods. Among the new genomic methods discussed in this edition are:

- Next-generation DNA sequencing, including the Illumina and 454 sequencing methods and platforms (Chapter 9)
- Applications of genomics, including the use of haplotypes to trace human migrations and phylogenetics to locate human genes associated with inherited diseases (Chapter 9)
- Forensic genotyping and the use of personalized genomics in medicine (Chapter 9)

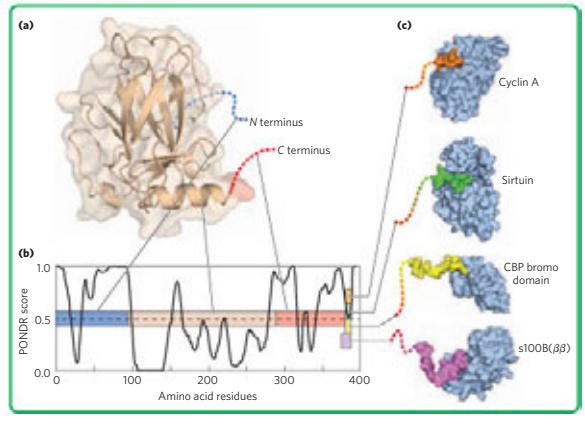


Next-generation reversible terminator sequencing

New Science

Every chapter has been thoroughly revised and updated to include both the most important advances in biochemistry and information needed in a modern biochemistry text. Among the new and updated topics in this edition are:

- Prebiotic evolution, black smokers, and the RNA world (Chapter 1)
- Intrinsically disordered proteins (Chapter 4)
- Transition-state analogs and irreversible inhibition (Chapter 6)
- Blood coagulation pathways in the context of enzymatic regulation (Chapter 6)



Binding of the intrinsically disordered carboxyl terminus of p53 to its binding partners

- Asymmetric lipid distribution in bilayers (Chapter 11)
- Role of BAR superfamily proteins in membrane curvature (Chapter 11)
- Scaffold proteins (AKAPS and others) and their regulatory roles (Chapter 12)
- Reactive oxygen species as byproducts and as signals (Chapter 19)
- Structure and function of the oxygen-evolving metal cluster in PSII (Chapter 19)
- Formation and transport of lipoproteins in mammals, including the roles of SREBP SCAP, and Insig in cholesterol regulation (Chapter 21)
- Integration of carbohydrate and lipid metabolism by PPARs, SREBPs, mTORC1, and LXR (Chapters 21, 23)
- Creatine phosphate and the role of creatine kinase in moving ATP to the cytosol (Chapter 23)
- Microbial symbionts in the gut and their influence on energy metabolism and adipogenesis (Chapter 23)
- Nucleosomes: their modification and positioning and higher-order chromatin structure (Chapter 24)
- DNA polymerases and homologous recombination (Chapter 25)
- Loading of eukaryotic RNA polymerase II (Chapter 26)
- Mutation-resistant nature of the genetic code (Chapter 27)
- Regulation of eukaryotic gene expression by miRNAs (Chapters 26 and 28).
- DNA looping, combinatorial control, chromatin remodeling, and positive regulation in eukaryotes (Chapter 28)
- Regulation of the initiation of transcription in eukaryotes (Chapter 28)
- Steroid-binding nuclear receptors (Chapter 28)

New Biochemical Methods

An appreciation of biochemistry often requires an understanding of how biochemical information is obtained. Some of the new methods or updates described in this edition are:

- Modern Sanger protein sequencing and mass spectrometry (Chapter 3)
- Mass spectrometry applied to proteomics, glycomics, lipidomics, and metabolomics (Chapters 3, 7, 10)
- Oligosaccharide microarrays to explore protein-oligosaccharide interactions and the "carbohydrate code" (Chapter 7)
- Modern genomic methods (Chapter 9)
- Genetic engineering of photosynthetic organisms (Chapter 20)
- Use of positron emission tomography (PET) to visualize tumors and brown adipose tissue (Chapter 23)
- Development of bacterial strains with altered genetic codes for site-specific insertion of novel amino acids into proteins (Chapter 27)

New Medical Applications

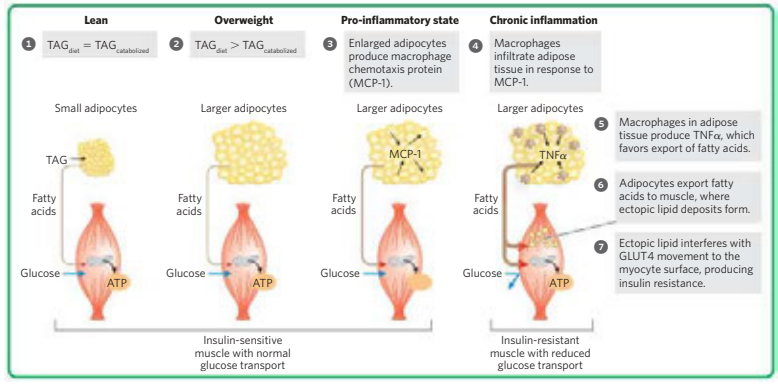
This icon is used throughout the book to denote material of special medical interest. As teachers, our goal is for students to learn biochemistry and to understand its relevance to a healthier life and a healthier planet. Many sections explore what we know about the molecular mechanisms of disease. A few of the new or revised medical applications in this edition are:

- Box 4-6, Death by Misfolding: The Prion Diseases
- Paganini and Ehlers-Danlos syndrome (Chapter 4)
- HIV protease inhibitors and how basic enzymatic principles influenced their design (Chapter 6)
- Blood coagulation cascade and hemophilia (Chapter 6)
- Curing African sleeping sickness with an enzymatic suicide inhibitor (Chapter 6)
- How researchers locate human genes involved in inherited diseases (Chapter 9)
- Multidrug resistance transporters and their importance in clinical medicine (Chapter 11)
- Multistep progression to colorectal cancer (Chapter 12)
- Cholesterol metabolism, cardiovascular disease, and mechanism of plaque formation in atherosclerosis (Chapter 21)
- P-450 and drug interactions (Chapter 21)
- HMG-CoA reductase (Chapter 21) and Box 21–3, The Lipid Hypothesis and the Development of Statins
- Box 24–1, Curing Disease by Inhibiting Topoisomerases, describing the use of topoisomerase inhibitors in the treatment of bacterial infections and cancer, including material on ciprofloxacin (the antibiotic effective for anthrax)
- Stem cells (Chapter 28)

Special Theme: Understanding Metabolism through Obesity and Diabetes

Obesity and its medical consequences—cardiovascular disease and diabetes—are fast becoming epidemic in the industrialized world, and we include new material on the biochemical connections between obesity and health throughout this edition. Our focus on diabetes provides an integrating theme throughout the chapters on metabolism and its control, and this will, we hope, inspire some students to find solutions for this disease. Some of the sections and boxes that highlight the interplay of metabolism, obesity, and diabetes are:

- Untreated Diabetes Produces Life-Threatening Acidosis (Chapter 2)
- Box 7–1, Blood Glucose Measurements in the Diagnosis and Treatment of Diabetes, introduces hemoglobin glycation and AGEs and their role in the pathology of advanced diabetes
- Glucose Uptake Is Deficient in Type 1 Diabetes Mellitus (Chapter 14)
- Ketone Bodies Are Overproduced in Diabetes and during Starvation (Chapter 17)
- Some Mutations in Mitochondrial Genomes Cause Disease (Chapter 19)
- Diabetes Can Result from Defects in the Mitochondria of Pancreatic β Cells (Chapter 19)
- Adipose Tissue Generates Glycerol 3-phosphate by Glyceroneogenesis (Chapter 21)
- Diabetes Mellitus Arises from Defects in Insulin Production or Action (Chapter 23)
- Section 23.4, Obesity and the Regulation of Body Mass, includes a new discussion of the roles of TORC1 in regulating cell growth
- Section 23.5, Obesity, the Metabolic Syndrome, and Type 2 Diabetes, discusses the role of ectopic lipids and inflammation in the development of insulin resistance and the management of type 2 diabetes with exercise, diet, and medication



Overloading adipocytes with triacylglycerols triggers inflammation in fat tissue, ectopic lipid deposition, and insulin resistance.

Special Theme: Evolution

Every time a biochemist studies a developmental pathway in nematodes, identifies key parts of an enzyme active site by determining what parts are conserved between species, or searches for the gene underlying a human genetic disease, he or she is relying on evolutionary theory. Funding agencies support the work in nematodes knowing that the insights will be relevant to humans. The conservation of functional residues in an enzyme active site telegraphs the shared history of every organism on the planet. More often than not, the search for a disease gene is a sophisticated exercise in phylogenetics. Evolution is thus a foundational concept to our discipline. Some of the many sections and boxes that deal with evolution include:

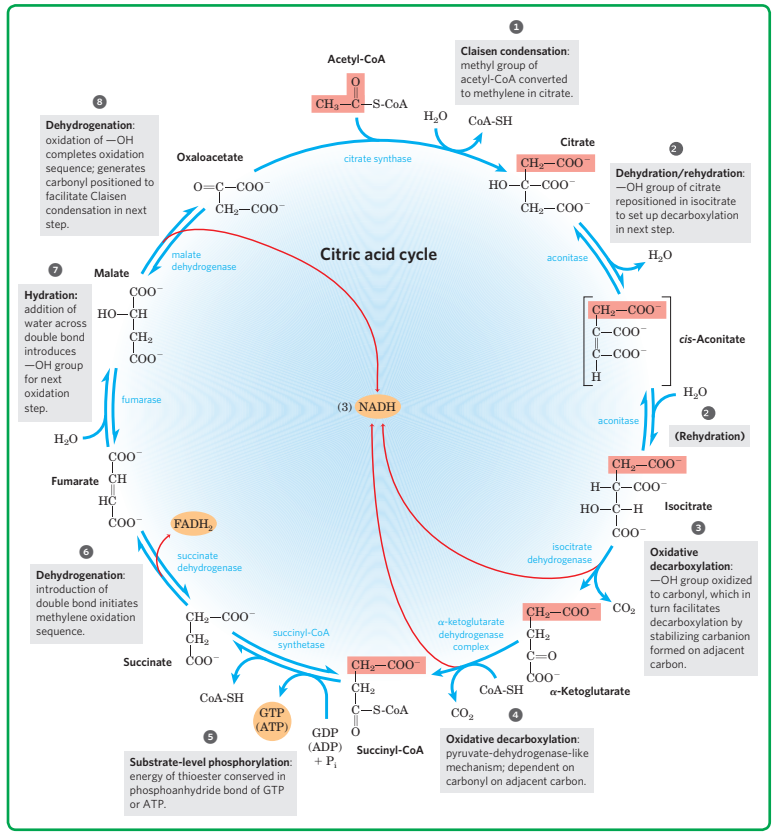
- Section 1.5, Evolutionary Foundations, discusses how life may have evolved and recounts some of the early milestones in the evolution of eukaryotic cells
- Genome Sequencing Informs Us about Our Humanity (Chapter 9)
- Genome Comparisons Help Locate Genes Involved in Disease (Chapter 9)
- Genome Sequences Inform Us about Our Past and Provide Opportunities for the Future (Chapter 9)
- Box 9–3, Getting to Know the Neanderthals
- ABC Transporters Use ATP to Drive the Active Transport of a Wide Variety of Substrates (Chapter 11)
- Signaling Systems of Plants Have Some of the Same Components Used by Microbes and Mammals (Chapter 12)
- The β-Oxidation Enzymes of Different Organelles Have Diverged during Evolution (Chapter 17)
- Section 19.10, The Evolution of Oxygenic Photosynthesis
- Mitochondria and Chloroplasts Evolved from Endosymbiotic Bacteria (Chapter 19)
- Photosystems I and II Evolved from Bacterial Photosystems (Chapter 19)
- RNA Synthesis Offers Important Clues to Biochemical Evolution (Chapter 26)
- Box 27–1, Exceptions That Prove the Rule: Natural Variations in the Genetic Code
- Box 27–2, From an RNA World to a Protein World
- Box 28-1, Of Fins, Wings, Beaks, and Things

Lehninger Teaching Hallmarks

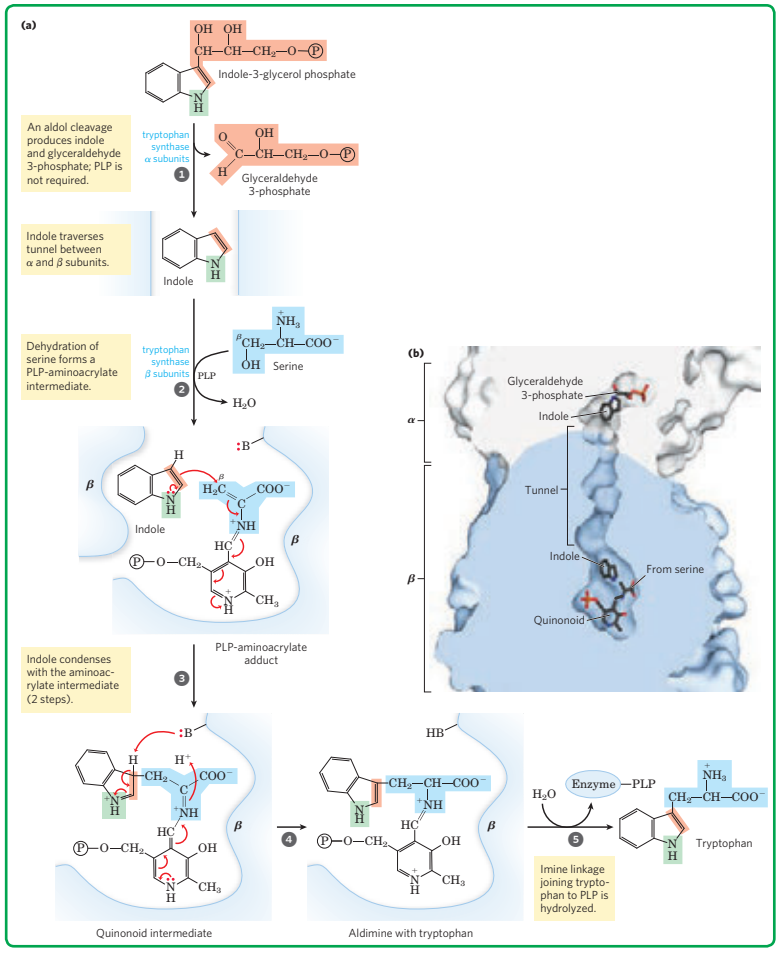
Students encountering biochemistry for the first time often have difficulty with two key aspects of the course: approaching quantitative problems and drawing on what they learned in organic chemistry to help them understand biochemistry. Those same students must also learn a complex language, with conventions that are often unstated. To help students cope with these challenges, we provide the following study aids:

Focus on Chemical Logic

- Section 13.2, **Chemical Logic and Common Biochemical Reactions**, discusses the common biochemical reaction types that underlie all metabolic reactions, helping students to connect organic chemistry with biochemistry.
- **NEW chemical logic figures** highlight the conservation of mechanism and illustrate patterns that make learning pathways easier. Chemical logic figures are provided for each of the central metabolic pathways, including glycolysis (Fig. 14–3), citric acid cycle (Fig. 16–7), and fatty acid oxidation (Fig. 17–9).



Reactions of the citric acid cycle



Tryptophan synthase reaction

- **Mechanism figures** feature step-by-step descriptions to help students understand the reaction process. These figures use a consistent set of conventions introduced and explained in detail with the first enzyme mechanism encountered (chymotrypsin, pp. 216–217).

Problem-Solving Tools

- **In-text Worked Examples** help students improve their quantitative problem-solving skills, taking them through some of the most difficult equations. New worked examples appear in Chapters 1, 2, and 19.
- **More than 600 end-of-chapter problems (about 25 of them new)** give students further opportunity to practice what they have learned.
- **Data Analysis Problems** (one at the end of each chapter), contributed by Brian White of the University of Massachusetts–Boston, encourage students to synthesize what they have learned and apply their knowledge to the interpretation of data from the literature.

Key Conventions

Many of the conventions that are so necessary for understanding each biochemical topic and the biochemical literature are broken out of the text and highlighted. These **Key Conventions** include clear statements of many assumptions and conventions that students are often expected to assimilate without being told (for example, peptide sequences are written from amino- to carboxyl-terminal end, left to right; nucleotide sequences are written from 5′ to 3′ end, left to right).

WORKED EXAMPLE 19–2 Stoichiometry of ATP Production: Effect of c Ring Size

(a) If *bovine* mitochondria have 8 c subunits per c ring, what is the predicted ratio of ATP formed per NADH oxidized? (b) What is the predicted value for *yeast* mitochondria, with 10 c subunits? (c) What are the comparable values for electrons entering the respiratory chain from $FADH_2$?

Solution: (a) The question asks us to determine how many ATP are produced per NADH. This is another way of asking us to calculate the P/O ratio, or x in Equation 19–11. If the c ring has 8 c subunits, then one full rotation will transfer 8 protons to the matrix and produce 3 ATP molecules. But this synthesis also requires the transport of 3 P_i into the matrix, at a cost of 1 proton each, adding 3 more protons to the total number required. This brings the total cost to (11 protons)/(3 ATP) = 3.7 protons/ATP. The consensus value for the number of protons pumped out per pair of electrons transferred from NADH is 10 (see Fig. 19–19). So, oxidizing 1 NADH produces (10 protons)/(3.7 protons/ATP) = 2.7 ATP.

(b) If the c ring has 10 c subunits, then one full rotation will transfer 10 protons to the matrix and produce 3 ATP molecules. Adding in the 3 protons to transport the 3 P_i into the matrix brings the total cost to (13 protons)/(3 ATP) = 4.3 protons/ATP. Oxidizing 1 NADH produces (10 protons)/(4.3 protons/ATP) = 2.3 ATP.

(c) When electrons enter the respiratory chain from $FADH_2$ (at ubiquinone), only 6 protons are available to drive ATP synthesis. This changes the calculation for bovine mitochondria to (6 protons)/(3.7protons/ ATP) = 1.6 ATP per pair of electrons from $FADH_2$. For yeast mitochondria, the calculation is (6 protons)/(4.3 protons/ATP) = 1.4 ATP per pair of electrons from $FADH_2$.

These calculated values of x or the P/O ratio define a range that includes the experimental values of 2.5 ATP/NADH and 1.5 ATP/$FADH_2$, and we therefore use these values throughout this book.

Media and Supplements

A full package of media resources and supplements provides instructors and students with innovative tools to support a variety of teaching and learning approaches. All these resources are fully integrated with the style and goals of the sixth-edition textbook.

NEW BiochemPortal (courses.bfwpub.com/lehninger6e)

This comprehensive and robust online teaching and learning tool incorporates the e-Book, all instructor and student resources, instructor assignment and gradebook functionality, and a new LearningCurve quizzing tool.

- BiochemPortal includes the **e-Book**, with the full contents of the text, highlighting and note-taking tools, and links to important media assets (listed below).
- In addition to all **instructor resources** (listed below), BiochemPortal provides instructors with the **ability to assign** any resource, as well as e-Book readings, discussion board posts, and their own materials. A gradebook tracks all student scores and can be easily exported to Excel or a campus Course Management System.
- **New** BiochemPortal also includes **LearningCurve**, a self-paced adaptive quizzing tool. With questions tailored to students' target difficulty level and an engaging scoring system, LearningCurve encourages students to incorporate content from the text into their study routine and provides them with a study plan on completion.
- Students can access any of the **student resources** provided with the text (see below) through links in the e-Book or the handy Resources tab.

e-Book (ebooks.bfwpub.com/lehninger6e)

This online version of the textbook combines the contents of the printed book with electronic study tools and a full complement of student media specifically created to support the text. The e-Book also provides useful material for instructors.

- **e-Book study tools** include instant navigation to any section or page of the book, bookmarks, highlighting, note-taking, instant search for any term, pop-up key-term definitions, and a spoken glossary.
- The text-specific **student media**, fully integrated throughout the e-Book, include animated enzyme mechanisms, animated biochemical techniques, problem-solving videos, molecular structure tutorials in Jmol, Protein Data Bank IDs in Jmol, and Living Graphs (each described under "Student Resources" below).
- **Instructor features** include the ability to add notes or files to any page and to share these notes with students. Notes may include text, Web links, animations, or photos. Instructors can also assign the entire text or a custom version of the e-Book.

Instructor Resources

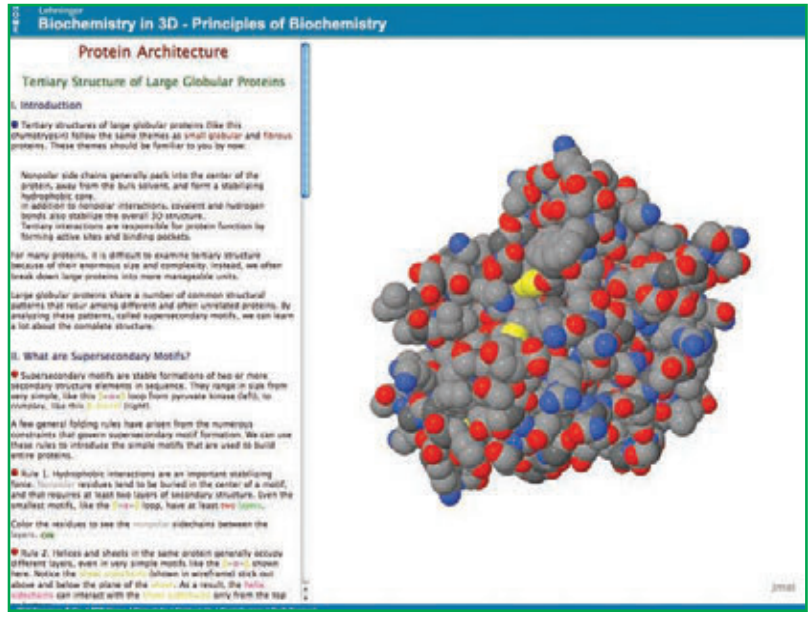
Instructors are provided with a comprehensive set of teaching tools, each developed to support the text, lecture presentations, and individual teaching styles. All instructor media are available for download on the **book website** (www.whfreeman.com/lehninger6e) and on the **Instructor Resource DVD** (ISBN 1-4641-0969-9).

- **New clicker questions** provide instructors with dynamic multiple-choice questions to be used with iClicker or other classroom response systems. The clicker questions have been written specifically to foster active learning in the classroom and better inform instructors on student misunderstandings.
- **Fully optimized JPEG files** of every figure, photo, and table in the text feature enhanced color, higher resolution, and enlarged fonts. The files have been reviewed by course instructors and tested in a large lecture hall to ensure maximum clarity and visibility. The JPEGs are also offered in separate files and in **PowerPoint** format for each chapter.
- **Animated Enzyme Mechanisms** and **Animated Biochemical Techniques** are available in Flash files and preloaded into PowerPoint, in both PC and Macintosh formats, for lecture presentation.
- A list of **Protein Data Bank IDs** for the structures in the text are arranged by figure number. A new feature in this edition is an index to all structures in the Jmol interactive Web browser applet.
- **Living Graphs**, illustrating key equations from the textbook, show the graphic results of changing parameters.
- A comprehensive **Test Bank** in PDF and editable Word formats includes 150 multiple-choice and short-answer problems per chapter, rated by level of difficulty.

Student Resources

Students are provided with media designed to enhance their understanding of biochemical principles and improve their problem-solving ability. All student media, along with the **PDB Structures** and **Living Graphs**, are also in the e-Book, and many are available on the book website (www.whfreeman.com/lehninger6e). Icons in the text indicate the availability of relevant animation, Living Graph, or Molecular Structure Tutorial.

- **Problem-Solving Videos**, created by Scott Ensign of Utah State University, provide 24/7 online problem-solving help to students. Through a two-part approach, each 10-minute video covers a key textbook problem representing a topic that students traditionally struggle to master. Dr. Ensign first describes a proven problem-solving strategy and then applies the strategy to the problem at hand in clear, concise steps. Students can easily pause, rewind, and review any steps as they wish until they firmly grasp not just the solution but also the reasoning behind it. Working through the problems in this way is designed to make students better and more confident at applying key strategies as they solve other textbook and exam problems.
- Student versions of the **Animated Enzyme Mechanisms** and **Animated Biochemical Techniques** help students understand key mechanisms and techniques at their own pace.



Protein Architecture Molecular Structure Tutorial

- **Molecular Structure Tutorials**, using the Jmol-Web browser applet, allow students to explore in more depth the molecular structures included in the textbook, including:

  Protein Architecture

  Bacteriorhodopsin

  Lac Repressor

  Nucleotides

  MHC Molecules

  Trimeric G Proteins

  Oxygen-Binding Proteins

  Restriction Endonucleases

  Hammerhead Ribozyme

The Absolute, Ultimate Guide to Lehninger Principles of Biochemistry, Sixth Edition, Study Guide and Solutions Manual, by Marcy Osgood (University of New Mexico School of Medicine) and Karen Ocorr (Sanford-Burnham Medical Research Institute); ISBN 1429294760

The Absolute, Ultimate Guide combines an innovative study guide with a reliable solutions manual (providing extended solutions to end-of-chapter problems) in one convenient volume. Thoroughly class-tested, the study guide includes for each chapter:

- **Major Concepts:** a road map through the chapter
- **What to Review:** questions that recap key points from previous chapters
- **Discussion Questions:** provided for each section; designed for individual review, study groups, or classroom discussion
- **A Self-Test:** "Do you know the terms?"; crossword puzzles; multiple-choice, fact-driven questions; and questions that ask students to apply their new knowledge in new directions—plus answers!

Acknowledgments

This book is a team effort, and producing it would be impossible without the outstanding people at W. H. Freeman and Company who supported us at every step along the way. Susan Moran (Senior Developmental Editor), Susan Winslow (Publisher), and Lauren Schultz (Senior Acquisitions Editor) helped develop the revision plan for this edition, made many helpful suggestions, encouraged us, and tried valiantly (if not always successfully) to keep us on schedule. Matthew Tontonoz (Developmental Editor) provided extremely helpful feedback on many chapters. Our outstanding Project Editor, Jane O'Neill, somehow kept the book moving through production in spite of our missed deadlines and last-minute changes, and did so with her usual grace and skill. We thank Art Director Diana Blume for her artistry in designing both the text and cover for the book. We appreciate the work of present and past copyeditors, including Karen Taschek, Liz Geller, and Linda Strange. Although Linda did not copyedit this edition, her lasting contributions from the first through the fifth editions are still clearly evident in the text. We thank Photo Research Manager Ted Szczepanski and Photo Researcher Elyse Rieder for their help in locating images and Courtney Lyons for help in orchestrating reviews and providing administrative assistance at many turns. We also thank Allison Michael, Media Editor, for assembling the ever more important media components to accompany the text. Our gratitude also goes to Debbie Clare, Associate Director of Marketing, for her creativity and good humor in coordinating the sales and marketing effort. A very special thanks goes to Kate Parker, who oversaw this project for the past three editions, and contributed much to its success, before moving on to other things; we will miss her insight, humor, and excellent taste in restaurants.

In Madison, Brook Soltvedt is (and has been for all the editions we have worked on) our first-line editor and critic. She is the first to see manuscript chapters, aids in manuscript and art development, ensures internal consistency in content and nomenclature, and keeps us on task with more-or-less gentle prodding. As she did for the fourth and fifth editions, Shelley Lusetti of New Mexico State University read every word of the text in proofs, caught numerous mistakes, and made many suggestions that improved the book.

The new art in this edition, including the new molecular graphics, was done by Adam Steinberg, here in Madison, who often made valuable suggestions that led to better and clearer illustrations. We feel very fortunate to have such gifted partners as Brook, Shelley, and Adam on our team.

We are also deeply indebted to Brian White of the University of Massachusetts–Boston, who wrote the data analysis problems at the end of each chapter.

Many colleagues played a special role through their interest in the project and their timely input. Prominent among these are Jeffrey D. Esko of the University of California, San Diego; and Jack Kirsch and his students at the University of California, Berkeley. Charles G. Hoogstraten of Michigan State University made many incisive and helpful comments on the manuscript and figures. We also thank Jeffrey A. Cohlberg of California State University at Long Beach for his critical comments. Many others helped us shape this sixth edition with their comments, suggestions, and criticisms. To all of them, we are deeply grateful:

Richard Amasino, *University of Wisconsin–Madison*
Laurens Anderson, *University of Wisconsin–Madison*
Alan Attie, *University of Wisconsin–Madison*
Kenneth Balazovich, *University of Michigan*
James Blankenship, *Cornell University*
Tracey Boncher, *Ferris State College of Pharmacy*
Brian Bothner, *Montana State University*
Mary Bryk, *Texas A&M University*
Sharada Buddha, *Saint Xavier University*
Jeff DeJong, *University of Texas, Dallas*
Keith Dunker, *Indiana University*
Kelly Elkins, *Metropolitan State College of Denver*
Gerald Feigenson, *Cornell University*
Brent Feske, *Armstrong Atlantic State University*

Marcello Forconi, *College of Charleston*
Wilson Francisco, *Arizona State University*
Greta Giles, *Georgia Gwinnett College*
Glenda Gillaspy, *Virginia Tech University*
Margaret Glasner, *Texas A&M University*
James Gober, *University of California, Los Angeles*
Burt Goldberg, *New York University*
Julie Gosse, *University of Maine*
Lesley Greene, *Old Dominion University*
Eric Hegg, *Michigan State University*
Justin Hines, *Lafayette College*
Peter Hinkle, *Cornell University*
Pui Ho, *Colorado State University*
David Hurley, *Gatton College of Pharmacy, ETSU*
Joseph Jez, *Washington University in St. Louis*
Kelly Johanson, *Xavier University of Louisiana*
Douglas Julin, *University of Maryland*
Mark Kearley, *Florida State University*
Dmitry Kolpashchikov, *University of Central Florida*
Min-Hao Kuo, *Michigan State University*
Nicole LaRonde-LeBlanc, *University of Maryland*
Scott Lefler, *Arizona State University*
Andy LiWang, *University of California, Merced*
Thomas Marsh, *University of St. Thomas*
Michele McGuirl, *The University of Montana*
Michael Mendenhall, *University of Kentucky*
David Merkler, *University of South Florida*
Debra Moriarity, *University of Alabama: Huntsville*
Hunter Moseley, *University of Louisville*
Allen Nicholson, *Temple University*
James Ntambi, *University of Wisconsin–Madison*
Neil Osheroff, *Vanderbilt University School of Medicine*
Donald Ourth, *University of Memphis*
Terry Platt, *University of Rochester*
Wendy Pogozelski, *State University of New York, Geneseo*
Joseph Provost, *Minnesota State University, Moorhead*
Gregory Raner, *University of North Carolina, Greensboro*
Lisa Rezende, *University of Arizona*
Douglas Root, *University of North Texas*
Johannes Rudolph, *University of Colorado*
Phillip Ryals, *University of West Florida*
Kevin Siebenlist, *Marquette University*
Kerry Smith, *Clemson University*
Julian Snow, *University of the Sciences*
Alejandra Stenger, *University of Illinois, Urbana–Champaign*
Amy Stockert, *Ohio Northern University*
Jon Stoltzfus, *Michigan State University*
Toni Vidal-Puig, *University of Cambridge*
Chuan Xiao, *University of Texas, El Paso*
Michael Yaffe, *Massachusetts Institute of Technology*
Laura Zapanta, *University of Pittsburgh*

We lack the space here to acknowledge all the other individuals whose special efforts went into this book. We offer instead our sincere thanks—and the finished book that they helped guide to completion. We, of course, assume full responsibility for errors of fact or emphasis.

We want especially to thank our students at the University of Wisconsin–Madison for their numerous comments and suggestions. If something in the book does not work, they are never shy about letting us know it. We are grateful to the students and staff of our research groups, who helped us balance the competing demands on our time; to our colleagues in the Department of Biochemistry at the University of Wisconsin–Madison, who helped us with advice and criticism; and to the many students and teachers who have written to suggest ways of improving the book. We hope our readers will continue to provide input for future editions.

Finally, we express our deepest appreciation to our wives, Brook and Beth, and our families, who showed extraordinary patience with, and support for, our book writing.

David L. Nelson
Michael M. Cox
Madison, Wisconsin
June 2012

Contents in Brief

Contents

1

The Foundations of Biochemistry

About fourteen billion years ago, the universe arose as a cataclysmic explosion of hot, energy-rich subatomic particles. Within seconds, the simplest elements (hydrogen and helium) were formed. As the universe expanded and cooled, material condensed under the influence of gravity to form stars. Some stars became enormous and then exploded as supernovae, releasing the energy needed to fuse simpler atomic nuclei into the more complex elements. Atoms and molecules formed swirling masses of dust particles, and their accumulation led eventually to the formation of rocks, planetoids, and planets. Thus were produced, over billions of years, Earth itself and the chemical elements found on Earth today. About four billion years ago, life arose—simple microorganisms with the ability to extract energy from chemical compounds and, later, from sunlight, which they used to make a vast array of more complex **biomolecules** from the simple elements and compounds on the Earth's surface. We and all other living organisms are made of stardust.

Biochemistry asks how the remarkable properties of living organisms arise from the thousands of different biomolecules. When these molecules are isolated and examined individually, they conform to all the physical and chemical laws that describe the behavior of inanimate matter—as do all the processes occurring in living organisms. The study of biochemistry shows how the collections of inanimate molecules that constitute living organisms interact to maintain and perpetuate life animated solely by the physical and chemical laws that govern the nonliving universe.

Yet organisms possess extraordinary attributes, properties that distinguish them from other collections of matter. What are these distinguishing features of living organisms?

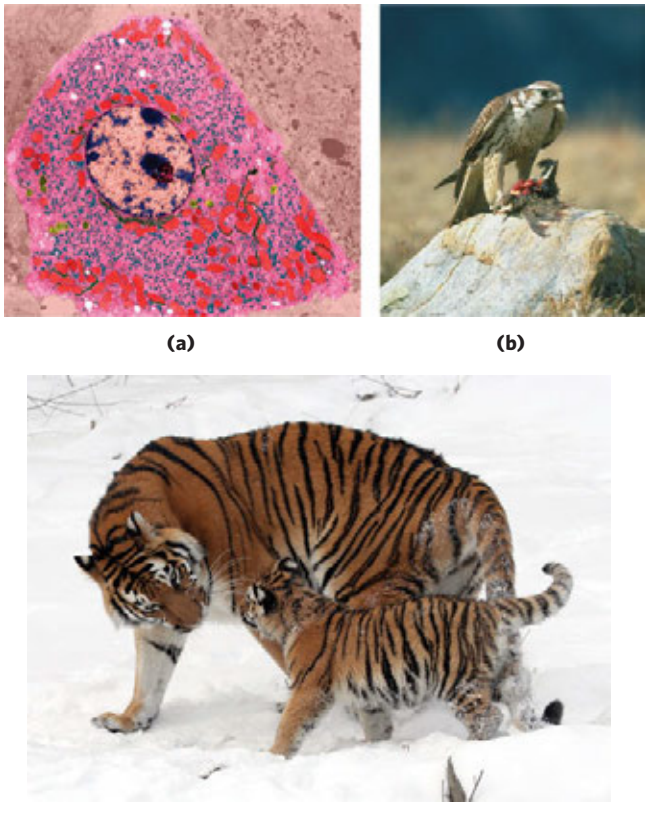
A high degree of chemical complexity and microscopic organization. Thousands of different molecules make up a cell's intricate internal structures **(Fig. 1–1a)**. These include very long polymers, each with its characteristic sequence of subunits, its unique three-dimensional structure, and its highly specific selection of binding partners in the cell.

Systems for extracting, transforming, and using energy from the environment (Fig. 1–1b), enabling organisms to build and maintain their intricate structures and to do mechanical, chemical, osmotic, and electrical work. This counteracts the tendency of all matter to decay toward a more disordered state, to come to equilibrium with its surroundings.

Defined functions for each of an organism's components and regulated interactions among them. This is true not only of macroscopic structures, such as leaves and stems or hearts and lungs, but also of microscopic intracellular structures and individual chemical compounds. The interplay among the chemical components of a living organism is dynamic; changes in one component cause coordinating or compensating changes in another, with the whole ensemble displaying a character beyond that of its individual parts. The collection of molecules carries out a program, the end result of which is reproduction of the program and self-perpetuation of that collection of molecules—in short, life.

Mechanisms for sensing and responding to alterations in their surroundings. Organisms constantly adjust to these changes by adapting their internal chemistry or their location in the environment.

A capacity for precise self-replication and self-assembly (Fig. 1–1c). A single bacterial cell placed in a sterile nutrient medium can give rise to



(a) (b) (c)

FIGURE 1–1 Some characteristics of living matter. (a) Microscopic complexity and organization are apparent in this colorized image of a thin section of a secretory cell from the pancreas, viewed with the electron microscope. **(b)** A prairie falcon acquires nutrients and energy by consuming a smaller bird. **(c)** Biological reproduction occurs with near-perfect fidelity.

a billion identical "daughter" cells in 24 hours. Each cell contains thousands of different molecules, some extremely complex; yet each bacterium is a faithful copy of the original, its construction directed entirely from information contained in the genetic material of the original cell. On a larger scale, the progeny of a vertebrate animal share a striking resemblance to their parents, also the result of their inheritance of parental genes.

A capacity to change over time by gradual evolution. Organisms change their inherited life strategies, in very small steps, to survive in new circumstances. The result of eons of evolution is an enormous diversity of life forms, superficially very different **(Fig. 1–2)** but fundamentally related through their shared ancestry. This fundamental unity of living organisms is reflected at the molecular level in the similarity of gene sequences and protein structures.

Despite these common properties, and the fundamental unity of life they reveal, it is difficult to make generalizations about living organisms. Earth has an enormous diversity of organisms. The range of habitats, from hot springs to Arctic tundra, from animal intestines

FIGURE 1–2 Diverse living organisms share common chemical features. Birds, beasts, plants, and soil microorganisms share with humans the same basic structural units (cells) and the same kinds of macromolecules (DNA, RNA, proteins) made up of the same kinds of monomeric subunits (nucleotides, amino acids). They utilize the same pathways for synthesis of cellular components, share the same genetic code, and derive from the same evolutionary ancestors. Shown here is a detail from ĐThe Garden of Eden,Đ by Jan van Kessel the Younger (1626Đ1679).

to college dormitories, is matched by a correspondingly wide range of specific biochemical adaptations, achieved within a common chemical framework. For the sake of clarity, in this book we sometimes risk certain generalizations, which, though not perfect, remain useful; we also frequently point out the exceptions to these generalizations, which can prove illuminating.

Biochemistry describes in molecular terms the structures, mechanisms, and chemical processes shared by all organisms and provides organizing principles that underlie life in all its diverse forms, principles we refer to collectively as *the molecular logic of life*. Although biochemistry provides important insights and practical applications in medicine, agriculture, nutrition, and industry, its ultimate concern is with the wonder of life itself.

In this introductory chapter we give an overview of the cellular, chemical, physical, and genetic backgrounds to biochemistry and the overarching principle of evolution—how life emerged and evolved into the diversity of organisms we see today. As you read through the book, you may find it helpful to refer back to this chapter at intervals to refresh your memory of this background material.

1.1 Cellular Foundations

The unity and diversity of organisms become apparent even at the cellular level. The smallest organisms consist of single cells and are microscopic. Larger, multicellular

organisms contain many different types of cells, which vary in size, shape, and specialized function. Despite these obvious differences, all cells of the simplest and most complex organisms share certain fundamental properties, which can be seen at the biochemical level.

Cells Are the Structural and Functional Units of All Living Organisms

Cells of all kinds share certain structural features **(Fig. 1–3)**. The **plasma membrane** defines the periphery of the cell, separating its contents from the surroundings. It is composed of lipid and protein molecules that form a thin, tough, pliable, hydrophobic barrier around the cell. The membrane is a barrier to the free passage of inorganic ions and most other charged or polar compounds. Transport proteins in the plasma membrane allow the passage of certain ions and molecules; receptor proteins transmit signals into the cell; and membrane enzymes participate in some reaction pathways. Because the individual lipids and proteins of the plasma membrane are not covalently linked, the entire structure is remarkably flexible, allowing changes in the shape and size of the cell. As a cell grows, newly made lipid and protein molecules are inserted into its plasma membrane; cell division produces two cells, each with its own membrane. This growth and cell division (fission) occurs without loss of membrane integrity.

The internal volume enclosed by the plasma membrane, the **cytoplasm** (Fig. 1–3), is composed of an aqueous solution, the **cytosol**, and a variety of suspended particles with specific functions. These particulate components (membranous organelles such as mitochondria and chloroplasts; supramolecular structures such as **ribosomes** and **proteasomes**, the sites of protein synthesis and degradation) sediment when cytoplasm is centrifuged at 150,000 *g* (*g* is the gravitational force of Earth). What remains as the supernatant fluid is the cytosol, a highly concentrated solution containing enzymes and the RNA molecules that encode them; the components (amino acids and nucleotides) from which these macromolecules are assembled; hundreds of small organic molecules called **metabolites**, intermediates in biosynthetic and degradative pathways; **coenzymes**, compounds essential to many enzyme-catalyzed reactions; and inorganic ions.

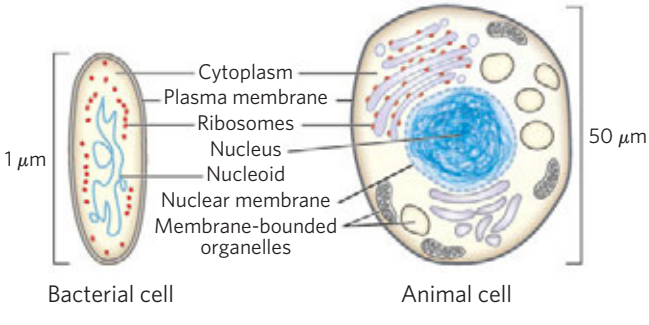


FIGURE 1–3 The universal features of living cells. All cells have a nucleus or nucleoid containing their DNA, a plasma membrane, and cytoplasm. The cytosol is defined as that portion of the cytoplasm that remains in the supernatant after gentle breakage of the plasma membrane and centrifugation of the resulting extract at 150,000 *g* for 1 hour. Eukaryotic cells contain a variety of membrane-bounded organelles (mitochondria, chloroplasts) and large particles (ribosomes, for example), which are sedimented by this centrifugation and can be recovered from the pellet.

All cells have, for at least some part of their life, either a **nucleoid** or a **nucleus**, in which the **genome**—the complete set of genes, composed of DNA—is replicated and stored, with its associated proteins. The nucleoid, in bacteria and archaea, is not separated from the cytoplasm by a membrane; the nucleus, in **eukaryotes**, is enclosed within a double membrane, the nuclear envelope. Cells with nuclear envelopes make up the large domain Eukarya (Greek *eu*, "true," and *karyon*, "nucleus"). Microorganisms without nuclear membranes, formerly grouped together as **prokaryotes** (Greek *pro*, "before"), are now recognized as comprising two very distinct groups: the domains Bacteria and Archaea, described below.

Cellular Dimensions Are Limited by Diffusion

Most cells are microscopic, invisible to the unaided eye. Animal and plant cells are typically 5 to 100 μm in diameter, and many unicellular microorganisms are only 1 to 2 μm long (see the inside back cover for information on units and their abbreviations). What limits the dimensions of a cell? The lower limit is probably set by the minimum number of each type of biomolecule required by the cell. The smallest cells, certain bacteria known as mycoplasmas, are 300 nm in diameter and have a volume of about 10^{-14} mL. A single bacterial ribosome is about 20 nm in its longest dimension, so a few ribosomes take up a substantial fraction of the volume in a mycoplasmal cell.

The upper limit of cell size is probably set by the rate of diffusion of solute molecules in aqueous systems. For example, a bacterial cell that depends on oxygen-consuming reactions for energy extraction must obtain molecular oxygen by diffusion from the surrounding medium through its plasma membrane. The cell is so small, and the ratio of its surface area to its volume is so large, that every part of its cytoplasm is easily reached by O_2 diffusing into the cell. With increasing cell size, however, surface-to-volume ratio decreases, until metabolism consumes O_2 faster than diffusion can supply it. Metabolism that requires O_2 thus becomes impossible as cell size increases beyond a certain point, placing a theoretical upper limit on the size of cells. Oxygen is only one of many low molecular weight species that must diffuse from outside the cell to various regions of its interior, and the same surface-to-volume argument applies to each of them as well.

There Are Three Distinct Domains of Life

All living organisms fall into one of three large groups (domains) that define three branches of the evolutionary tree of life originating from a common progenitor

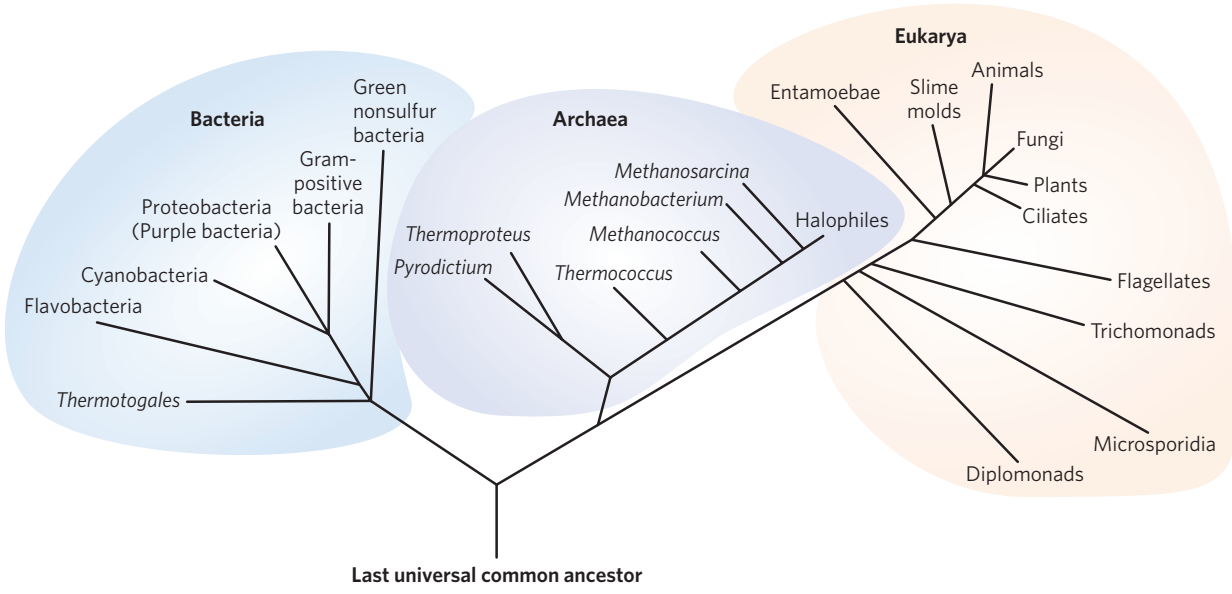


FIGURE 1–4 Phylogeny of the three domains of life. Phylogenetic relationships are often illustrated by a Ðfamily treeÐ of this type. The basis for this tree is the similarity in nucleotide sequences of the ribosomal RNAs of each group; the more similar the sequence, the closer the location of the branches, with the distance between branches representing the degree of difference between two sequences. Phylogenetic trees can also be constructed from similarities across species of the amino acid sequences of a single protein. For example, sequences of the protein GroEL (a bacterial protein that assists in protein folding) were compared to generate the tree in Figure 3Ð35. The tree in Figure 3Ð36 is a ÐconsensusÐ tree, which uses several comparisons such as these to derive the best estimates of evolutionary relatedness among a group of organisms.

(Fig. 1–4). Two large groups of single-celled microorganisms can be distinguished on genetic and biochemical grounds: **Bacteria** and **Archaea**. Bacteria inhabit soils, surface waters, and the tissues of other living or decaying organisms. Many of the Archaea, recognized as a distinct domain by Carl Woese in the 1980s, inhabit extreme environments—salt lakes, hot springs, highly acidic bogs, and the ocean depths. The available evidence suggests that the Archaea and Bacteria diverged early in evolution. All eukaryotic organisms, which make up the third domain, **Eukarya**, evolved from the same branch that gave rise to the Archaea; eukaryotes are therefore more closely related to archaea than to bacteria.

Within the domains of Archaea and Bacteria are subgroups distinguished by their habitats. In **aerobic** habitats with a plentiful supply of oxygen, some resident organisms derive energy from the transfer of electrons from fuel molecules to oxygen within the cell. Other environments are **anaerobic**, virtually devoid of oxygen, and microorganisms adapted to these environments obtain energy by transferring electrons to nitrate (forming N_2), sulfate (forming H_2S), or CO_2 (forming CH_4). Many organisms that have evolved in anaerobic environments are *obligate* anaerobes: they die when exposed to oxygen. Others are *facultative* anaerobes, able to live with or without oxygen.

Organisms Differ Widely in Their Sources of Energy and Biosynthetic Precursors

We can classify organisms according to how they obtain the energy and carbon they need for synthesizing cellular material (as summarized in **Fig. 1–5**). There are two broad categories based on energy sources: **phototrophs** (Greek *trophē*, "nourishment") trap and use sunlight, and **chemotrophs** derive their energy from oxidation of a chemical fuel. Some chemotrophs oxidize inorganic fuels—HS^- to S^0 (elemental sulfur), S^0 to to or to for example. Phototrophs and chemotrophs may be further divided into those that can synthesize all of their biomolecules directly from CO_2 (**autotrophs**) and those that require some preformed organic nutrients made by other organisms (**heterotrophs**). We can describe an organism's mode of nutrition by combining these terms. For example, cyanobacteria are photoautotrophs; humans are chemoheterotrophs. Even finer distinctions can be made, and many organisms can obtain energy and carbon from more than one source under different environmental or developmental conditions.

Bacterial and Archaeal Cells Share Common Features but Differ in Important Ways

The best-studied bacterium, *Escherichia coli*, is a usually harmless inhabitant of the human intestinal tract. The *E. coli* cell **(Fig. 1–6a)** is an ovoid about 2 μm long and a little less than 1 μm in diameter, but other bacteria may be spherical or rod-shaped. It has a protective outer membrane and an inner plasma membrane that encloses the cytoplasm and the nucleoid. Between the inner and outer membranes is a thin but strong layer of a high molecular weight polymer (peptidoglycan) that gives the cell its shape and rigidity. The plasma membrane and the layers outside it constitute

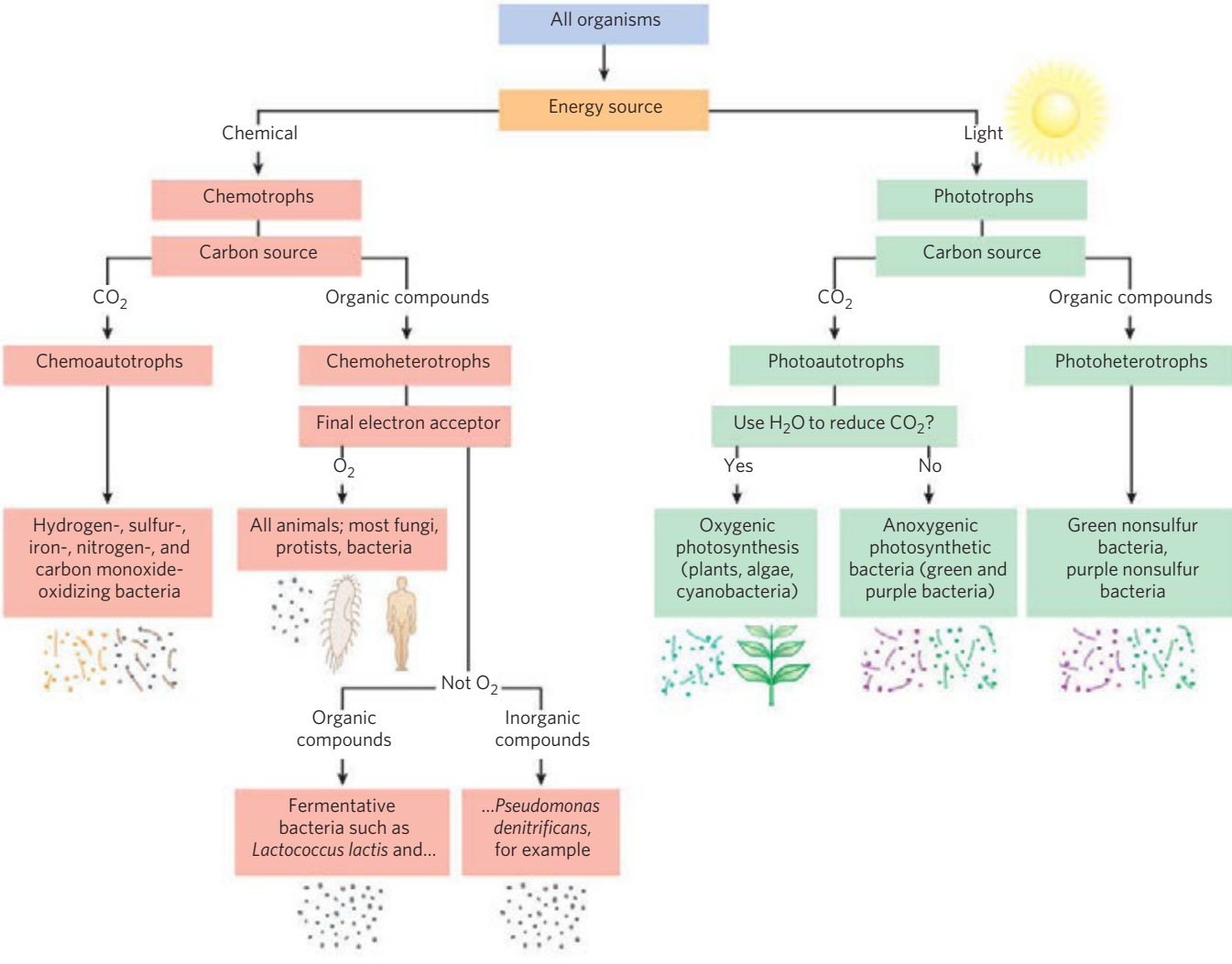


FIGURE 1–5 All organisms can be classified according to their source of energy (sunlight or oxidizable chemical compounds) and their source of carbon for the synthesis of cellular material.

the **cell envelope**. The plasma membranes of bacteria consist of a thin bilayer of lipid molecules penetrated by proteins. Archaeal plasma membranes have a similar architecture, but the lipids can be strikingly different from those of bacteria (see Fig. 10–12). Bacteria and archaea have group-specific specializations of their cell envelopes (Fig. 1–6b–d). Some bacteria, called gram-positive because they are colored by Gram's stain (introduced by Hans Peter Gram in 1882), have a thick layer of peptidoglycan outside their plasma membrane but lack an outer membrane. Gram-negative bacteria have an outer membrane composed of a lipid bilayer into which are inserted complex lipopolysaccharides and proteins called porins that provide transmembrane channels for low molecular weight compounds and ions to diffuse across this outer membrane. The structures outside the plasma membrane of archaea differ from organism to organism, but they, too, have a layer of peptidoglycan or protein that confers rigidity on their cell envelopes.

The cytoplasm of *E. coli* contains about 15,000 ribosomes, various numbers (10 to thousands) of copies of each of 1,000 or so different enzymes, perhaps 1,000 organic compounds of molecular weight less than 1,000 (metabolites and cofactors), and a variety of inorganic ions. The nucleoid contains a single, circular molecule of DNA, and the cytoplasm (like that of most bacteria) contains one or more smaller, circular segments of DNA called **plasmids**. In nature, some plasmids confer resistance to toxins and antibiotics in the environment. In the laboratory, these DNA segments are especially amenable to experimental manipulation and are powerful tools for genetic engineering (see Chapter 9).

Other species of bacteria, as well as archaea, contain a similar collection of biomolecules, but each species has physical and metabolic specializations related to its environmental niche and nutritional sources. Cyanobacteria, for example, have internal membranes specialized to trap energy from light (see Fig. 19–67). Many archaea live in extreme environments and have biochemical

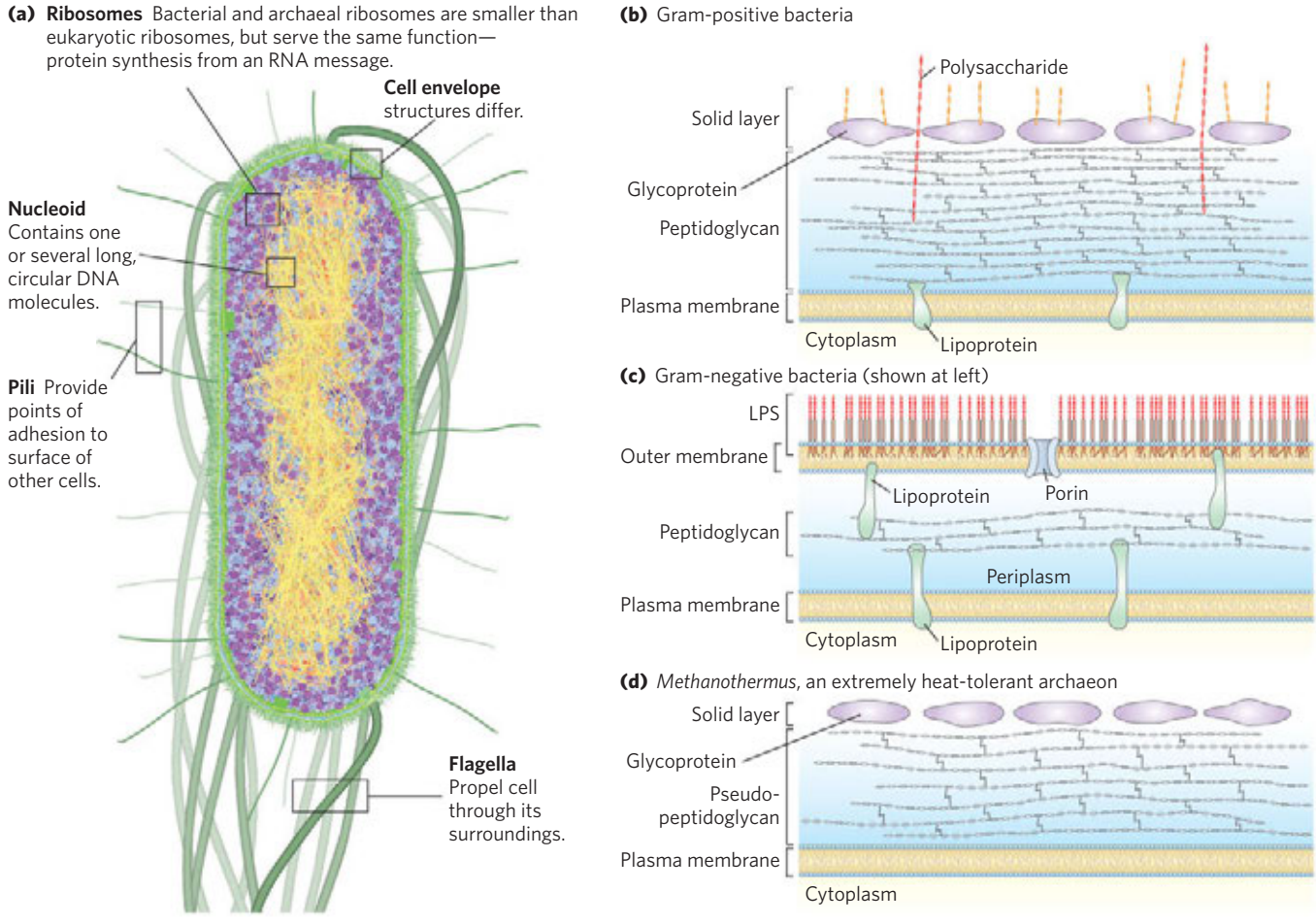


FIGURE 1–6 Some common structural features of bacterial and archaeal cells. (a) This correct-scale drawing of *E. coli* serves to illustrate some common features. **(b)** The cell envelope of gram-positive bacteria is a single membrane with a thick, rigid layer of peptidoglycan on its outside surface. A variety of polysaccharides and other complex polymers are interwoven with the peptidoglycan, and surrounding the whole is a porous Ðsolid layerÐ composed of glycoproteins. **(c)** *E. coli* is gram-negative and has a double membrane. Its outer membrane has a lipopolysaccharide (LPS) on the outer surface and phospholipids on the inner surface. This outer membrane is studded with protein channels (porins) that allow small molecules, but not proteins, to diffuse through. The inner (plasma) membrane, made of phospholipids and proteins, is impermeant to both large and small molecules. Between the inner and outer membranes, in the periplasm, is a thin layer of peptidoglycan, which gives the cell shape and rigidity, but does not retain Gram's stain. **(d)** Archaeal membranes vary in structure and composition, but all have a single membrane surrounded by an outer layer that includes either a peptidoglycanlike structure, a porous protein shell (solid layer), or both.

adaptations to survive in extremes of temperature, pressure, or salt concentration. Differences in ribosomal structure gave the first hints that Bacteria and Archaea constituted separate domains. Most bacteria (including *E. coli*) exist as individual cells, but often associate in biofilms or mats, in which large numbers of cells adhere to each other and to some solid substrate beneath or at an aqueous surface. Cells of some bacterial species (the myxobacteria, for example) show simple social behavior, forming many-celled aggregates in response to signals between neighboring cells.

Eukaryotic Cells Have a Variety of Membranous Organelles, Which Can Be Isolated for Study

Typical eukaryotic cells **(Fig. 1–7)** are much larger than bacteria—commonly 5 to 100 μm in diameter, with cell volumes a thousand to a million times larger than those of bacteria. The distinguishing characteristics of eukaryotes are the nucleus and a variety of membrane-enclosed organelles with specific functions. These organelles include **mitochondria**, the site of most of the energy-extracting reactions of the cell; the **endoplasmic reticulum** and **Golgi complexes**, which play central roles in the synthesis and processing of lipids and membrane proteins; **peroxisomes**, in which very long-chain fatty acids are oxidized; and **lysosomes**, filled with digestive enzymes to degrade unneeded cellular debris. In addition to these, plant cells also contain **vacuoles** (which store large quantities of organic acids) and **chloroplasts** (in which sunlight drives the synthesis of ATP in the process of photosynthesis) (Fig. 1–7). Also present in the cytoplasm of many cells are granules or droplets containing stored nutrients such as starch and fat.

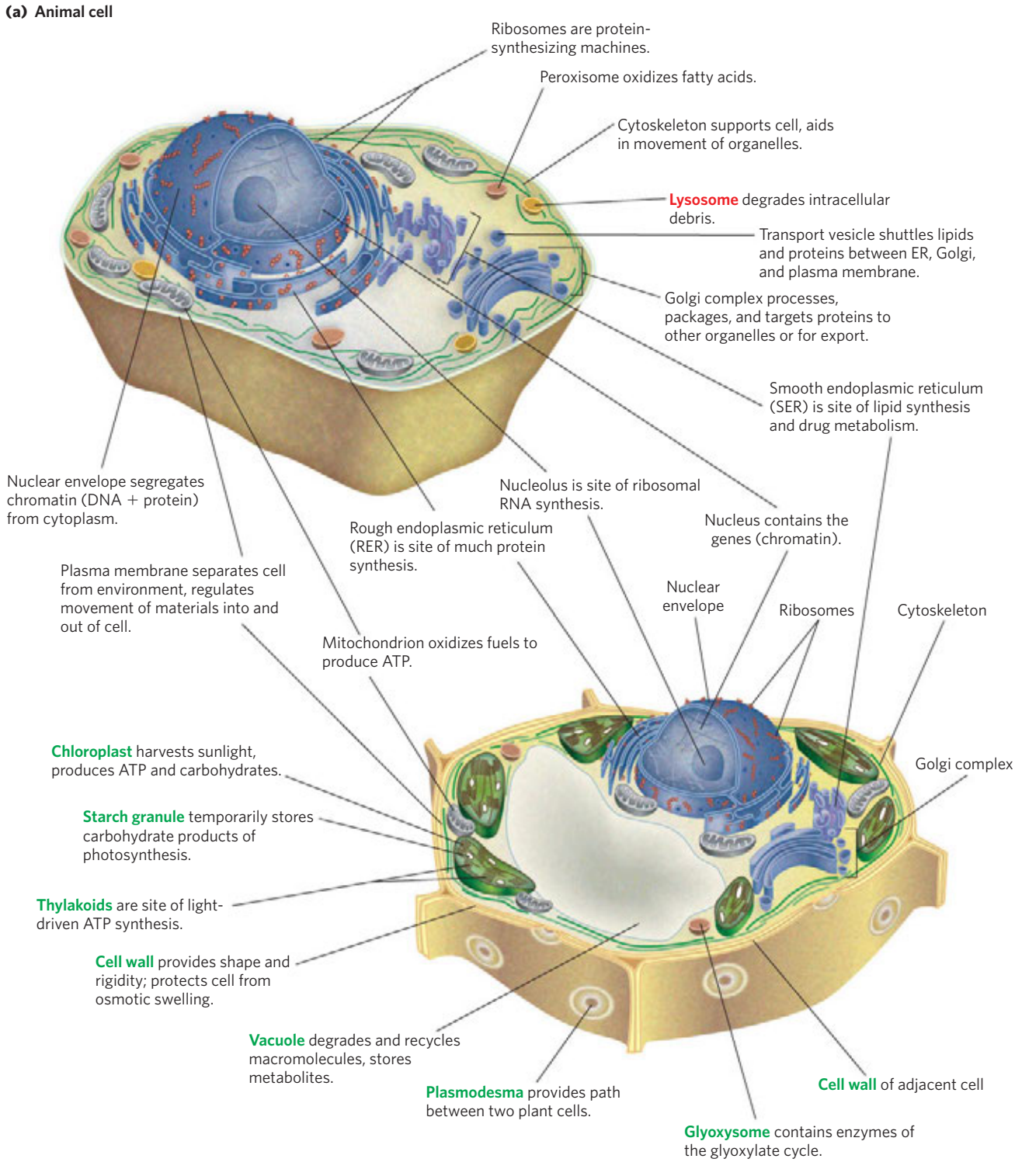


FIGURE 1–7 Eukaryotic cell structure. Schematic illustrations of two major types of eukaryotic cell: **(a)** a representative animal cell and **(b)** a representative plant cell. Plant cells are usually 10 to 100 μm in diameterÐ larger than animal cells, which typically range from 5 to 30 μm. Structures labeled in red are unique to animal cells; those labeled in green are unique to plant cells. Eukaryotic microorganisms (such as protists and fungi) have structures similar to those in plant and animal cells, but many also contain specialized organelles not illustrated here.

In a major advance in biochemistry, Albert Claude, Christian de Duve, and George Palade developed methods for separating organelles from the cytosol and from each other—an essential step in investigating their structures and functions. In a typical cell fractionation **(Fig. 1–8)**, cells or tissues in solution are gently disrupted by physical shear. This treatment ruptures the plasma membrane but leaves most of the organelles intact. The homogenate is then centrifuged; organelles such as nuclei, mitochondria, and lysosomes differ in size and therefore sediment at different rates.

These methods were used to establish, for example, that lysosomes contain degradative enzymes, mitochondria contain oxidative enzymes, and chloroplasts contain photosynthetic pigments. The isolation of an organelle enriched in a certain enzyme is often the first step in the purification of that enzyme.

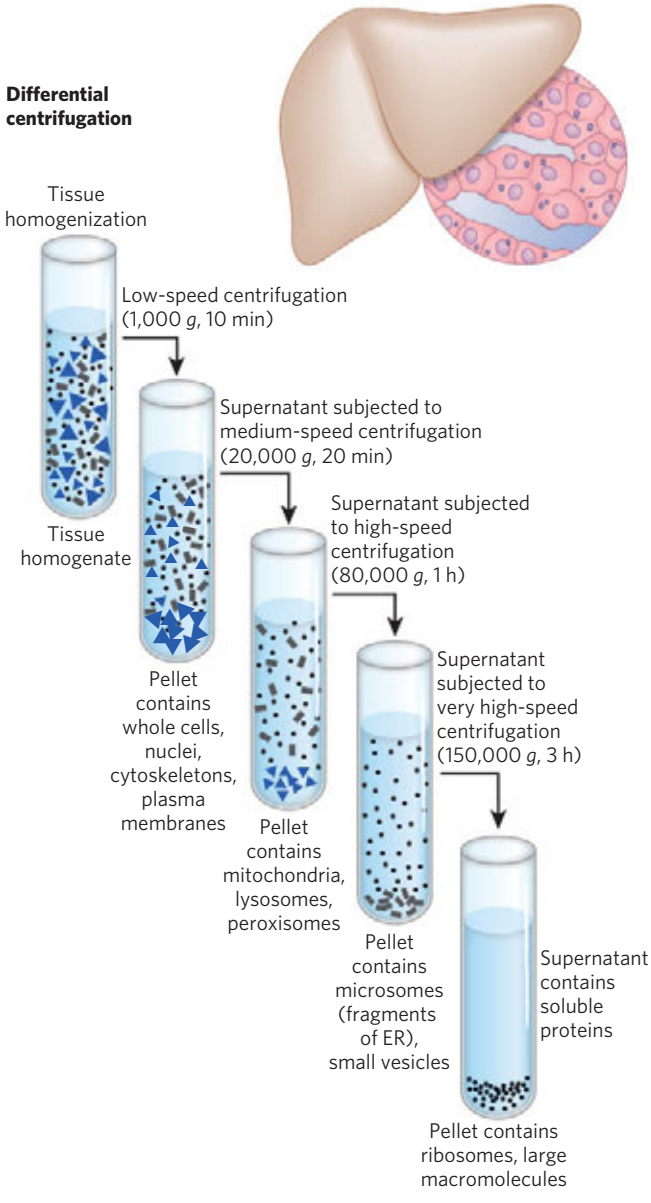


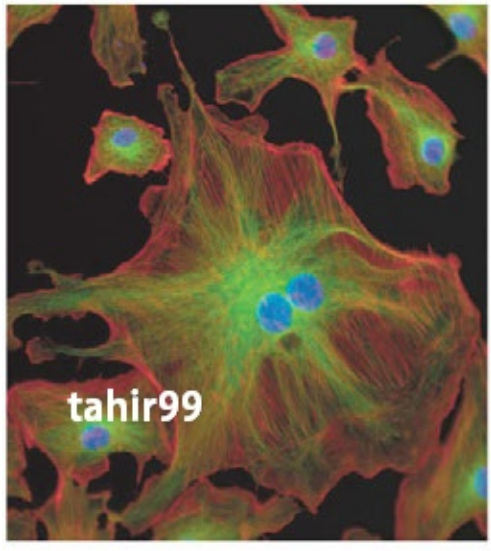
FIGURE 1–8 Subcellular fractionation of tissue. A tissue such as liver is first mechanically homogenized to break cells and disperse their contents in an aqueous buffer. The sucrose medium has an osmotic pressure similar to that in organelles, thus balancing diffusion of water into and out of the organelles, which would swell and burst in a solution of lower osmolarity (see Fig. 2Ð13). The large and small particles in the suspension can be separated by centrifugation at different speeds. Larger particles sediment more rapidly than small particles, and soluble material does not sediment. By careful choice of the conditions of centrifugation, subcellular fractions can be separated for biochemical characterization.

The Cytoplasm Is Organized by the Cytoskeleton and Is Highly Dynamic

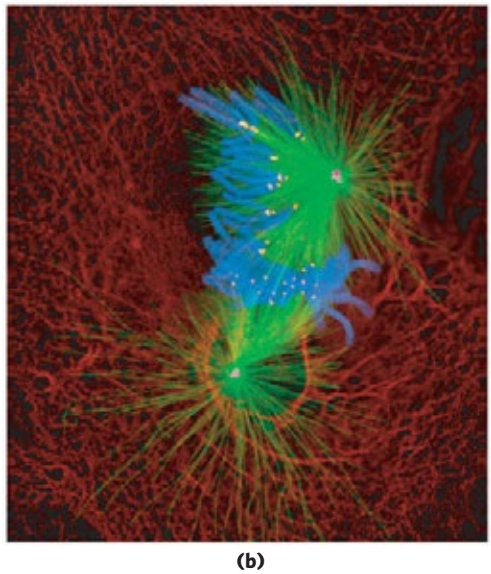
Fluorescence microscopy reveals several types of protein filaments crisscrossing the eukaryotic cell, forming an interlocking three-dimensional meshwork, the **cytoskeleton**. There are three general types of cytoplasmic filaments—actin filaments, microtubules, and intermediate filaments **(Fig. 1–9)**—differing in width (from about 6 to 22 nm), composition, and specific function. All types provide structure and organization to the cytoplasm and shape to the cell. Actin filaments and microtubules also help to produce the motion of organelles or of the whole cell.

Each type of cytoskeletal component is composed of simple protein subunits that associate noncovalently to form filaments of uniform thickness. These filaments are not permanent structures; they undergo constant disassembly into their protein subunits and reassembly into filaments. Their locations in cells are not rigidly fixed but may change dramatically with mitosis, cytokinesis, amoeboid motion, or changes in cell shape. The assembly, disassembly, and location of all types of filaments are regulated by other proteins, which serve to link or bundle the filaments or to move cytoplasmic organelles along the filaments. (Bacteria contain actinlike proteins that serve similar roles in those cells.)

The picture that emerges from this brief survey of eukaryotic cell structure is of a cell with a meshwork of structural fibers and a complex system of membrane-enclosed compartments (Fig. 1–7). The filaments disassemble and then reassemble elsewhere. Membranous vesicles bud from one organelle and fuse with another. Organelles move through the cytoplasm along protein filaments, their motion powered by energy-dependent motor proteins. The **endomembrane system** segregates specific metabolic processes and provides surfaces on which certain enzyme-catalyzed reactions occur. **Exocytosis** and **endocytosis**, mechanisms of transport (out of and into cells, respectively) that involve membrane fusion and fission, provide paths between the cytoplasm and surrounding medium, allowing for secretion of substances produced in the cell and uptake of extracellular materials.



(a)



(b)

FIGURE 1–9 The three types of cytoskeletal filaments: actin filaments, microtubules, and intermediate filaments. Cellular structures can be labeled with an antibody (that recognizes a characteristic protein) covalently attached to a fluorescent compound. The stained structures are visible when the cell is viewed with a fluorescence microscope. **(a)** Endothelial cells from the bovine pulmonary artery. Bundles of actin filaments called Ðstress fibersÐ are stained red; microtubules, radiating from the cell center, are stained green; and chromosomes (in the nucleus) are stained blue. **(b)** A newt lung cell undergoing mitosis. Microtubules (green), attached to structures called kinetochores (yellow) on the condensed chromosomes (blue), pull the chromosomes to opposite poles, or centrosomes (magenta), of the cell. Intermediate filaments, made of keratin (red), maintain the structure of the cell.

Although complex, this organization of the cytoplasm is far from random. The motion and positioning of organelles and cytoskeletal elements are under tight regulation, and at certain stages in its life, a eukaryotic cell undergoes dramatic, finely orchestrated reorganizations, such as the events of mitosis. The interactions between the cytoskeleton and organelles are noncovalent, reversible, and subject to regulation in response to various intracellular and extracellular signals.

Cells Build Supramolecular Structures

Macromolecules and their monomeric subunits differ greatly in size **(Fig. 1–10)**. An alanine molecule is less than 0.5 nm long. A molecule of hemoglobin, the oxygen-carrying protein of erythrocytes (red blood cells), consists of nearly 600 amino acid subunits in four long chains, folded into globular shapes and associated in a structure 5.5 nm in diameter. In turn, proteins are much smaller than ribosomes (about 20 nm in diameter), which are in turn much smaller than organelles such as mitochondria, typically 1,000 nm in diameter. It is a long jump from simple biomolecules to cellular structures that can be seen with the light microscope. **Figure 1–11** illustrates the structural hierarchy in cellular organization.

The monomeric subunits of proteins, nucleic acids, and polysaccharides are joined by covalent bonds. In supramolecular complexes, however, macromolecules are held together by noncovalent interactions—much weaker, individually, than covalent bonds. Among these noncovalent interactions are hydrogen bonds (between polar groups), ionic interactions (between charged groups), hydrophobic interactions (among nonpolar groups in aqueous solution), and van der Waals interactions (London forces)—all of which have energies much smaller than those of covalent bonds. These noncovalent interactions are described in Chapter 2. The large numbers of weak interactions between macromolecules in supramolecular complexes stabilize these assemblies, producing their unique structures.

In Vitro Studies May Overlook Important Interactions among Molecules

One approach to understanding a biological process is to study purified molecules in vitro ("in glass"—in the test tube), without interference from other molecules present in the intact cell—that is, in vivo ("in the living"). Although this approach has been remarkably revealing, we must keep in mind that the inside of a cell is quite different from the inside of a test tube. The "interfering" components eliminated by purification may be critical to the biological function or regulation of the molecule purified. For example, in vitro studies of pure enzymes are commonly done at very low enzyme concentrations in thoroughly stirred aqueous solutions. In the cell, an enzyme is dissolved or suspended in the gel-like cytosol with thousands of other proteins, some of which bind to that enzyme and influence its activity. Some enzymes are components of multienzyme complexes in which reactants are channeled from one enzyme to another, never entering the bulk solvent. When all of the known macromolecules in a cell are

represented at their known dimensions and concentrations **(Fig. 1–12)**, it is clear that the cytosol is very crowded and that diffusion of macromolecules within the cytosol must be slowed by collisions with other large structures. In short, a given molecule may behave quite differently in the cell and in vitro. A central challenge of biochemistry is to understand the influences of cellular organization and macromolecular associations on the function of individual enzymes and other biomolecules—to understand function in vivo as well as in vitro.

SUMMARY 1.1 Cellular Foundations

- All cells are bounded by a plasma membrane; have a cytosol containing metabolites, coenzymes, inorganic ions, and enzymes; and have a set of genes contained within a nucleoid (bacteria and archaea) or nucleus (eukaryotes).
- All organisms require a source of energy to perform cellular work. Phototrophs obtain energy from sunlight; chemotrophs oxidize chemical fuels, passing electrons to good electron acceptors: inorganic compounds, organic compounds, or molecular oxygen.
- Bacterial and archaeal cells contain cytosol, a nucleoid, and plasmids, all contained within a cell envelope. Eukaryotic cells have a nucleus and are multicompartmented, with certain processes

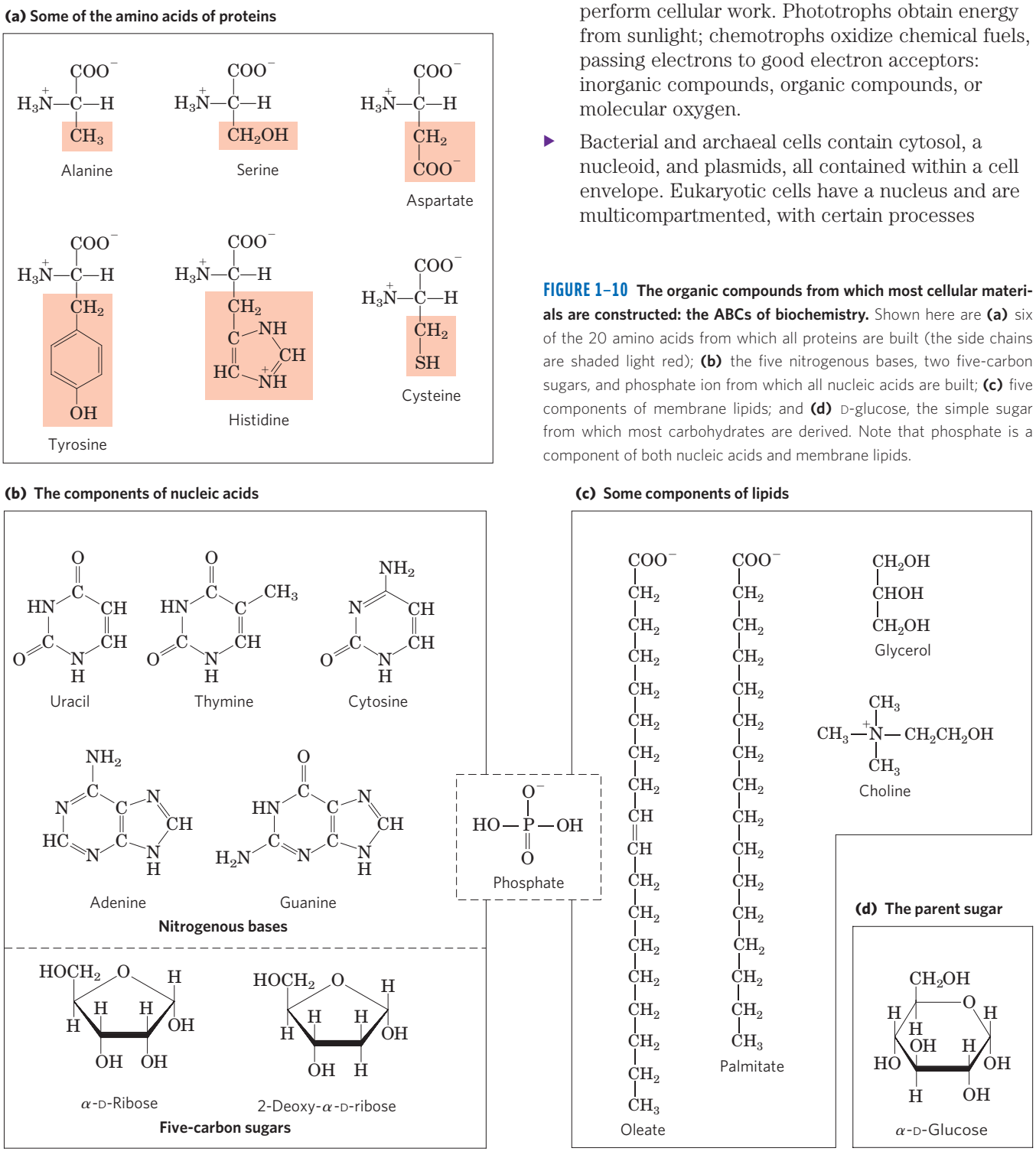


FIGURE 1–10 The organic compounds from which most cellular materials are constructed: the ABCs of biochemistry. Shown here are **(a)** six of the 20 amino acids from which all proteins are built (the side chains are shaded light red); **(b)** the five nitrogenous bases, two five-carbon sugars, and phosphate ion from which all nucleic acids are built; **(c)** five components of membrane lipids; and **(d)** D-glucose, the simple sugar from which most carbohydrates are derived. Note that phosphate is a component of both nucleic acids and membrane lipids.

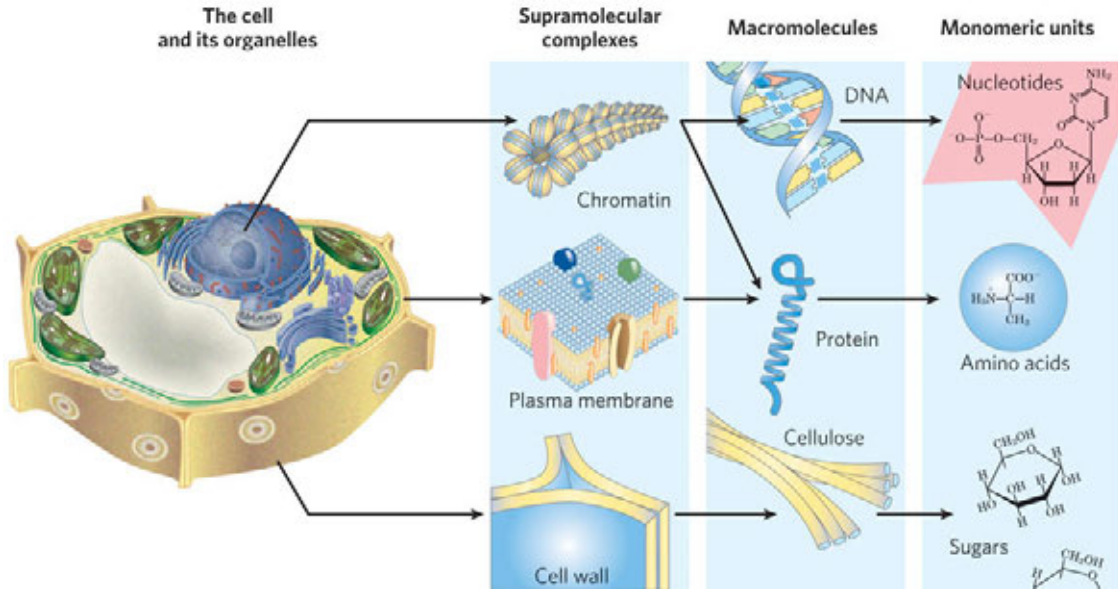


FIGURE 1–11 Structural hierarchy in the molecular organization of cells. The organelles and other relatively large components of cells are composed of supramolecular complexes, which in turn are composed of smaller macromolecules and even smaller molecular subunits. For example, the nucleus of this plant cell contains chromatin, a supramolecular complex that consists of DNA and basic proteins (histones). DNA is made up of simple monomeric subunits (nucleotides), as are proteins (amino acids).

segregated in specific organelles; organelles can be separated and studied in isolation.

- Cytoskeletal proteins assemble into long filaments that give cells shape and rigidity and serve as rails along which cellular organelles move throughout the cell.
- Supramolecular complexes held together by noncovalent interactions are part of a hierarchy of structures, some visible with the light microscope. When individual molecules are removed from these complexes to be studied in vitro, interactions important in the living cell may be lost.

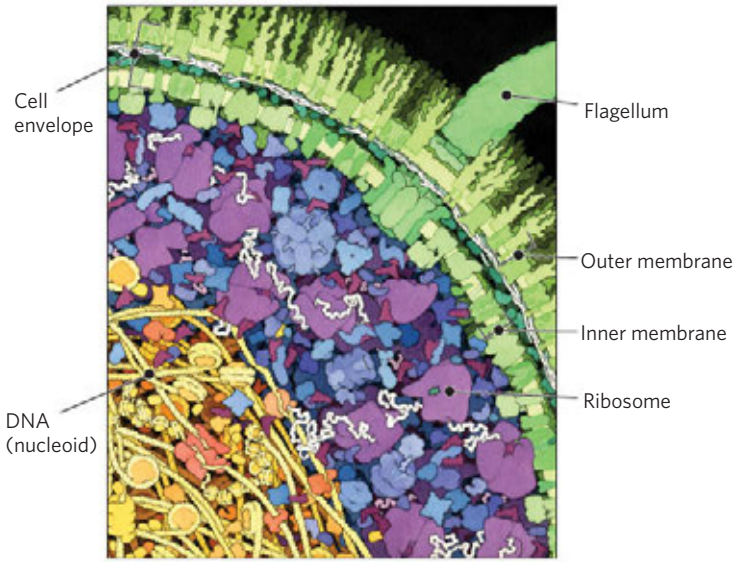


FIGURE 1–12 The crowded cell. This drawing by David Goodsell is an accurate representation of the relative sizes and numbers of macromolecules in one small region of an *E. coli* cell. This concentrated cytosol, crowded with proteins and nucleic acids, is very different from the typical extract of cells used in biochemical studies, in which the cytosol has been diluted manyfold and the interactions between diffusing macromolecules have been strongly altered.

1.2 Chemical Foundations

Biochemistry aims to explain biological form and function in chemical terms. By the late eighteenth century, chemists had concluded that the composition of living matter is strikingly different from that of the inanimate world. Antoine-Laurent Lavoisier (1743–1794) noted the relative chemical simplicity of the "mineral world" and contrasted it with the complexity of the "plant and animal worlds"; the latter, he knew, were composed of compounds rich in the elements carbon, oxygen, nitrogen, and phosphorus.

During the first half of the twentieth century, parallel biochemical investigations of glucose breakdown in yeast and in animal muscle cells revealed remarkable chemical similarities in these two apparently very different cell types; the breakdown of glucose in yeast and muscle cells involved the same 10 chemical intermediates, and the same 10 enzymes. Subsequent studies of many other biochemical processes in many different organisms have confirmed the generality of this observation, neatly summarized in 1954 by Jacques Monod: "What is true of *E. coli* is true of the elephant." The current understanding that all organisms share a common evolutionary origin is based in part on this observed

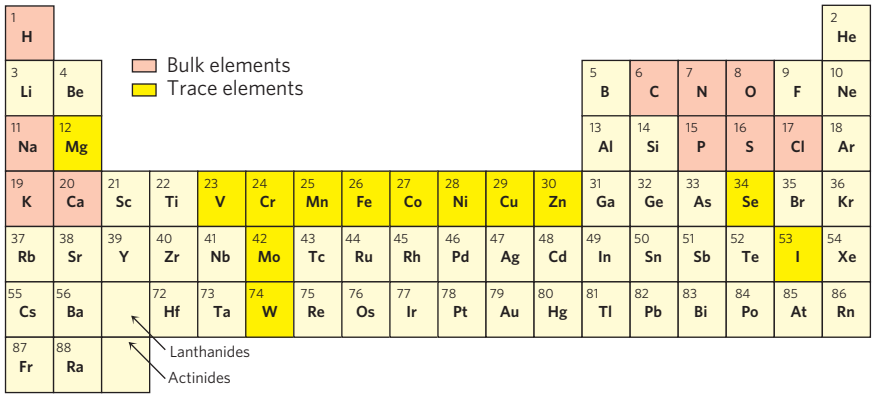


FIGURE 1–13 Elements essential to animal life and health. Bulk elements (shaded light red) are structural components of cells and tissues and are required in the diet in gram quantities daily. For trace elements (shaded yellow), the requirements are much smaller: for humans, a few milligrams per day of Fe, Cu, and Zn, even less of the others. The elemental requirements for plants and microorganisms are similar to those shown here; the ways in which they acquire these elements vary.

universality of chemical intermediates and transformations, often termed "biochemical unity."

Fewer than 30 of the more than 90 naturally occurring chemical elements are essential to organisms. Most of the elements in living matter have relatively low atomic numbers; only three have atomic numbers above that of selenium, 34 **(Fig. 1–13)**. The four most abundant elements in living organisms, in terms of percentage of total number of atoms, are hydrogen, oxygen, nitrogen, and carbon, which together make up more than 99% of the mass of most cells. They are the lightest elements capable of efficiently forming one, two, three, and four bonds, respectively; in general, the lightest elements form the strongest bonds. The trace elements (Fig. 1–13) represent a miniscule fraction of the weight of the human body, but all are essential to life, usually because they are essential to the function of specific proteins, including many enzymes. The oxygen-transporting capacity of the hemoglobin molecule, for example, is absolutely dependent on four iron ions that make up only 0.3% of its mass.

Biomolecules Are Compounds of Carbon with a Variety of Functional Groups

The chemistry of living organisms is organized around carbon, which accounts for more than half the dry weight of cells. Carbon can form single bonds with hydrogen atoms, and both single and double bonds with oxygen and nitrogen atoms **(Fig. 1–14)**. Of greatest significance in biology is the ability of carbon atoms to form very stable single bonds with up to four other carbon atoms. Two carbon atoms also can share two (or three) electron pairs, thus forming double (or triple) bonds.

The four single bonds that can be formed by a carbon atom project from the nucleus to the four apices of a tetrahedron **(Fig. 1–15)**, with an angle of about 109.5° between any two bonds and an average bond length of 0.154 nm. There is free rotation around each single bond, unless very large or highly charged groups are attached to both carbon atoms, in which case rotation may be restricted. A double bond is shorter (about 0.134 nm) and rigid, and allows only limited rotation about its axis.

Covalently linked carbon atoms in biomolecules can form linear chains, branched chains, and cyclic structures. It seems likely that the bonding versatility of carbon, with itself and with other elements, was a major factor in the selection of carbon compounds for the molecular machinery of cells during the origin and evolution of living organisms. No other chemical element can form molecules of such widely different sizes, shapes, and composition.

Most biomolecules can be regarded as derivatives of hydrocarbons, with hydrogen atoms replaced by a variety of functional groups that confer specific chemical properties on the molecule, forming various families of organic compounds. Typical of these are alcohols, which have one or more hydroxyl groups; amines, with amino groups; aldehydes and ketones, with carbonyl groups; and carboxylic acids, with carboxyl groups **(Fig. 1–16)**. Many biomolecules are polyfunctional, containing two or more types of functional groups

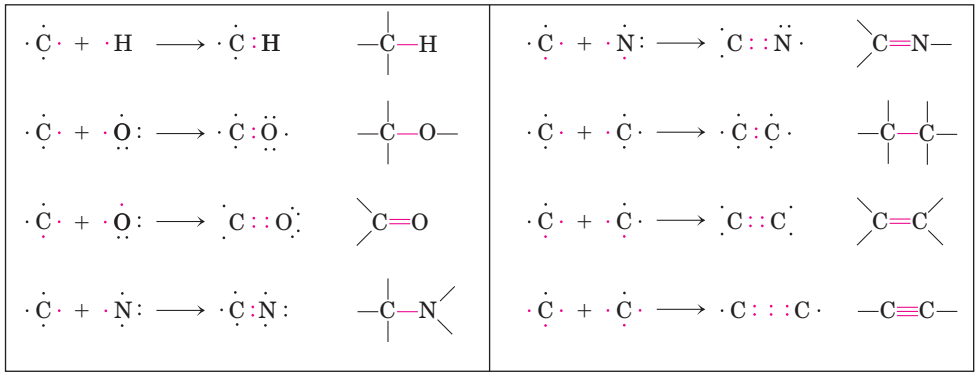


FIGURE 1–14 Versatility of carbon bonding. Carbon can form covalent single, double, and triple bonds (all bonds in red), particularly with other carbon atoms. Triple bonds are rare in biomolecules.

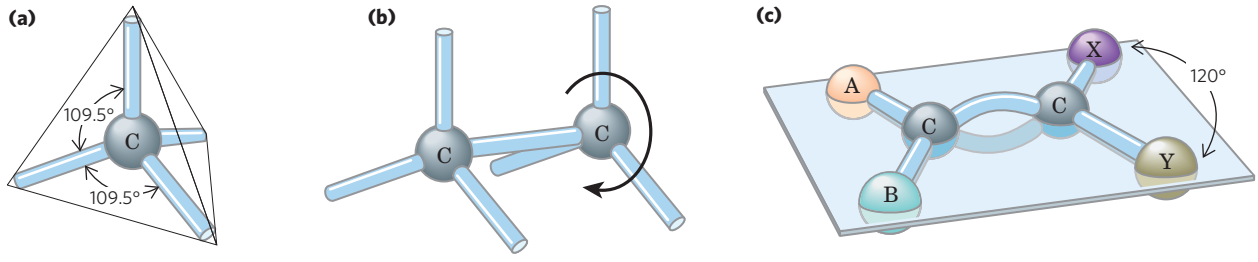


FIGURE 1–15 Geometry of carbon bonding. (a) Carbon atoms have a characteristic tetrahedral arrangement of their four single bonds. **(b)** Carbon–carbon single bonds have freedom of rotation, as shown for the compound ethane (CH_3—CH_3). **(c)** Double bonds are shorter and do not allow free rotation. The two doubly bonded carbons and the atoms designated A, B, X, and Y all lie in the same rigid plane.

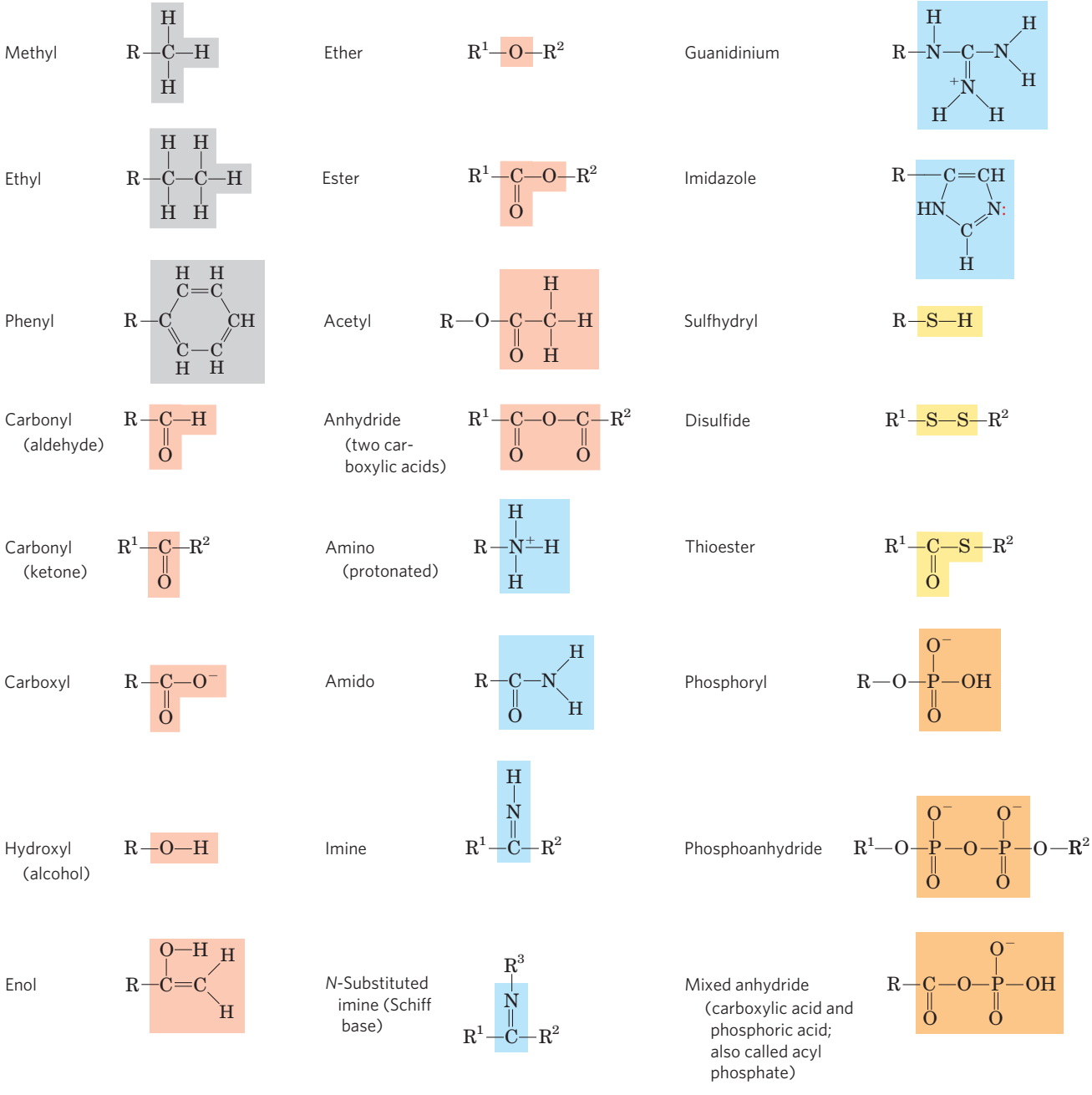


FIGURE 1–16 Some common functional groups of biomolecules. Functional groups are screened with a color typically used to represent the element that characterizes the group: gray for C, red for O, blue for N, yellow for S, and orange for P. In this figure and throughout the book, we use R to represent "any substituent." It may be as simple as a hydrogen atom, but typically it is a carbon-containing group. When two or more substituents are shown in a molecule, we designate them R^1, R^2, and so forth.

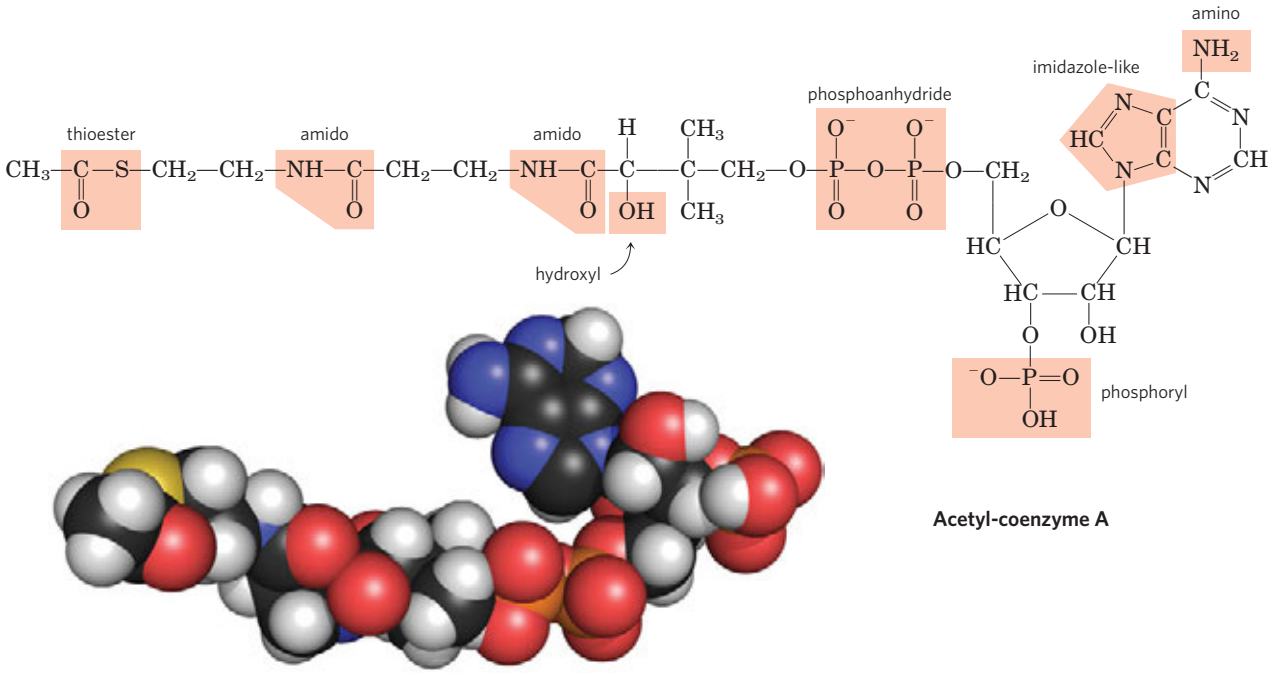


FIGURE 1–17 Several common functional groups in a single biomolecule. Acetyl-coenzyme A (often abbreviated as acetyl-CoA) is a carrier of acetyl groups in some enzymatic reactions. Its functional groups are screened in the structural formula. As we will see in Chapter 2, several of these functional groups can exist in protonated or unprotonated forms, depending on the pH. In the space-filling model, N is blue, C is black, P is orange, O is red, and H is white. The yellow atom at the left is the sulfur of the critical thioester bond between the acetyl moiety and coenzyme A.

(Fig. 1–17), each with its own chemical characteristics and reactions. The chemical "personality" of a compound is determined by the chemistry of its functional groups and their disposition in three-dimensional space.

Cells Contain a Universal Set of Small Molecules

Dissolved in the aqueous phase (cytosol) of all cells is a collection of perhaps a thousand different small organic molecules (M_r ~100 to ~500), with intracellular concentrations ranging from nanomolar to millimolar (see Fig. 15–4). (See Box 1–1 for an explanation of the various ways of referring to molecular weight.) These are the central metabolites in the major pathways occurring in nearly every cell—the metabolites and pathways that have been conserved throughout the course of evolution. This collection of molecules includes the common amino acids, nucleotides, sugars and their phosphorylated derivatives, and mono-, di-, and tricarboxylic acids. The molecules may be polar or charged, and are water-soluble. They are trapped in the cell because the plasma membrane is impermeable to them, although

BOX 1–1 Molecular Weight, Molecular Mass, and Their Correct Units

There are two common (and equivalent) ways to describe molecular mass; both are used in this text. The first is *molecular weight*, or *relative molecular mass*, denoted M_r. The molecular weight of a substance is defined as the ratio of the mass of a molecule of that substance to one-twelfth the mass of carbon-12 (^{12}C). Since M_r is a ratio, it is dimensionless—it has no associated units. The second is *molecular mass*, denoted m. This is simply the mass of one molecule, or the molar mass divided by Avogadro's number. The molecular mass, m, is expressed in daltons (abbreviated Da). One dalton is equivalent to one-twelfth the mass of carbon-12; a kilodalton (kDa) is 1,000 daltons; a megadalton (MDa) is 1 million daltons.

Consider, for example, a molecule with a mass 1,000 times that of water. We can say of this molecule either M_r = 18,000 or m = 18,000 daltons. We can also describe it as an "18 kDa molecule." However, the expression M_r = 18,000 daltons is incorrect.

Another convenient unit for describing the mass of a single atom or molecule is the atomic mass unit (formerly amu, now commonly denoted u). One atomic mass unit (1 u) is defined as one-twelfth the mass of an atom of carbon-12. Since the experimentally measured mass of an atom of carbon-12 is 1.9926×10^{-23} g, 1 u = 1.6606×10^{-24} g. The atomic mass unit is convenient for describing the mass of a peak observed by mass spectrometry (see Chapter 3, p. 100).

specific membrane transporters can catalyze the movement of some molecules into and out of the cell or between compartments in eukaryotic cells. The universal occurrence of the same set of compounds in living cells reflects the evolutionary conservation of metabolic pathways that developed in the earliest cells.

There are other small biomolecules, specific to certain types of cells or organisms. For example, vascular plants contain, in addition to the universal set, small molecules called **secondary metabolites**, which play roles specific to plant life. These metabolites include compounds that give plants their characteristic scents and colors, and compounds such as morphine, quinine, nicotine, and caffeine that are valued for their physiological effects on humans but used for other purposes by plants.

The entire collection of small molecules in a given cell under a specific set of conditions has been called the **metabolome**, in parallel with the term "genome." **Metabolomics** is the systematic characterization of the metabolome under very specific conditions (such as following administration of a drug or a biological signal such as insulin).

Macromolecules Are the Major Constituents of Cells

Many biological molecules are **macromolecules**, polymers with molecular weights above ~5,000 that are assembled from relatively simple precursors. Shorter polymers are called **oligomers** (Greek *oligos*, "few"). Proteins, nucleic acids, and polysaccharides are macromolecules composed of monomers with molecular weights of 500 or less. Synthesis of macromolecules is a major energy-consuming activity of cells. Macromolecules themselves may be further assembled into supramolecular complexes, forming functional units such as ribosomes. Table 1–1 shows the major classes of biomolecules in an *E. coli* cell.

TABLE 1–1 Molecular Components of an *E. coli* Cell

	Percentage of total weight of cell	Approximate number of different molecular species
Water	70	1
Proteins	15	3,000
Nucleic acids		
DNA	1	1–4
RNA	6	>3,000
Polysaccharides	3	10
Lipids	2	20
Monomeric subunits and intermediates	2	500
Inorganic ions	1	20

Proteins, long polymers of amino acids, constitute the largest fraction (besides water) of a cell. Some proteins have catalytic activity and function as enzymes; others serve as structural elements, signal receptors, or transporters that carry specific substances into or out of cells. Proteins are perhaps the most versatile of all biomolecules; a catalog of their many functions would be very long. The sum of all the proteins functioning in a given cell is the cell's **proteome**, and **proteomics** is the systematic characterization of this protein complement under a specific set of conditions. The **nucleic acids**, DNA and RNA, are polymers of nucleotides. They store and transmit genetic information, and some RNA molecules have structural and catalytic roles in supramolecular complexes. The **genome** is the entire sequence of a cell's DNA (or in the case of RNA viruses, its RNA), and **genomics** is the characterization of the comparative structure, function, evolution, and mapping of genomes. The **polysaccharides**, polymers of simple sugars such as glucose, have three major functions: as energy-rich fuel stores, as rigid structural components of cell walls (in plants and bacteria), and as extracellular recognition elements that bind to proteins on other cells. Shorter polymers of sugars (oligosaccharides) attached to proteins or lipids at the cell surface serve as specific cellular signals. A cell's **glycome** is all its carbohydrate-containing molecules. The **lipids**, water-insoluble hydrocarbon derivatives, serve as structural components of membranes, energy-rich fuel stores, pigments, and intracellular signals. The lipid-containing molecules in a cell constitute its **lipidome**. With the application of sensitive methods with great resolving power (mass spectrometry, for example), it is possible to distinguish and quantify hundreds or thousands of these components, and therefore to quantify their variations in response to changing conditions, signals, or drugs. Systems biology is an approach that tries to integrate the information from genomics, proteomics, glycomics, and lipidomics to give a molecular picture of all the activities of a cell under a given set of conditions, and the changes that occur when the system is perturbed by external signals or circumstances or by mutations.

Proteins, polynucleotides, and polysaccharides have large numbers of monomeric subunits and thus high molecular weights—in the range of 5,000 to more than 1 million for proteins, up to several *billion* for nucleic acids, and in the millions for polysaccharides such as starch. Individual lipid molecules are much smaller (M_r 750 to 1,500) and are not classified as macromolecules. But they can associate noncovalently into very large structures. Cellular membranes are built of enormous noncovalent aggregates of lipid and protein molecules.

Given their characteristic information-rich subunit sequences, proteins and nucleic acids are often referred to as **informational macromolecules**. Some oligosaccharides, as noted above, also serve as informational molecules.

Three-Dimensional Structure Is Described by Configuration and Conformation

The covalent bonds and functional groups of a biomolecule are, of course, central to its function, but so also is the arrangement of the molecule's constituent atoms in three-dimensional space—its stereochemistry. Carbon-containing compounds commonly exist as **stereoisomers**, molecules with the same chemical bonds and same chemical formula but different **configuration**, the fixed spatial arrangement of atoms. Interactions between biomolecules are invariably **stereospecific**, requiring specific configurations in the interacting molecules.

Figure 1–18 shows three ways to illustrate the stereochemistry, or configuration, of simple molecules. The perspective diagram specifies stereochemistry unambiguously, but bond angles and center-to-center bond lengths are better represented with ball-and-stick models. In space-filling models, the radius of each "atom" is proportional to its van der Waals radius, and the contours of the model define the space occupied by the molecule (the volume of space from which atoms of other molecules are excluded).

Configuration is conferred by the presence of either (1) double bonds, around which there is little or no freedom of rotation, or (2) chiral centers, around which

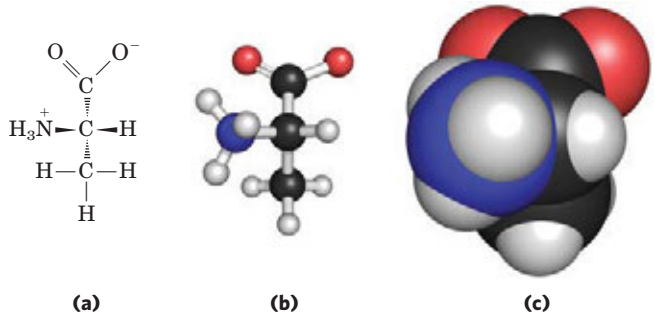
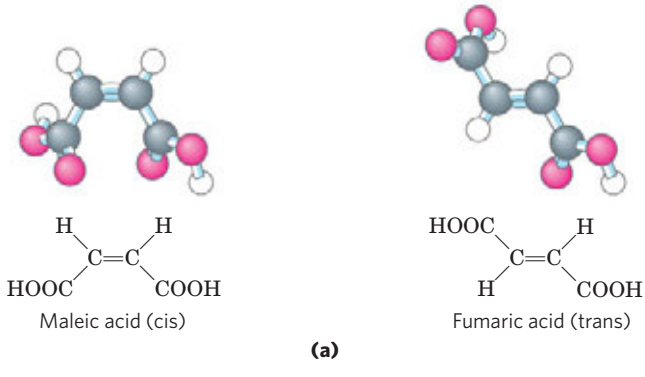


FIGURE 1–18 Representations of molecules. Three ways to represent the structure of the amino acid alanine (shown here in the ionic form found at neutral pH). **(a)** Structural formula in perspective form: a solid wedge (—) represents a bond in which the atom at the wide end projects out of the plane of the paper, toward the reader; a dashed wedge (┄) represents a bond extending behind the plane of the paper. **(b)** Ball-and-stick model, showing bond angles and relative bond lengths. **(c)** Space-filling model, in which each atom is shown with its correct relative van der Waals radius.

substituent groups are arranged in a specific orientation. The identifying characteristic of stereoisomers is that they cannot be interconverted without temporarily breaking one or more covalent bonds. **Figure 1–19a** shows the configurations of maleic acid and its isomer, fumaric acid. These compounds are **geometric**



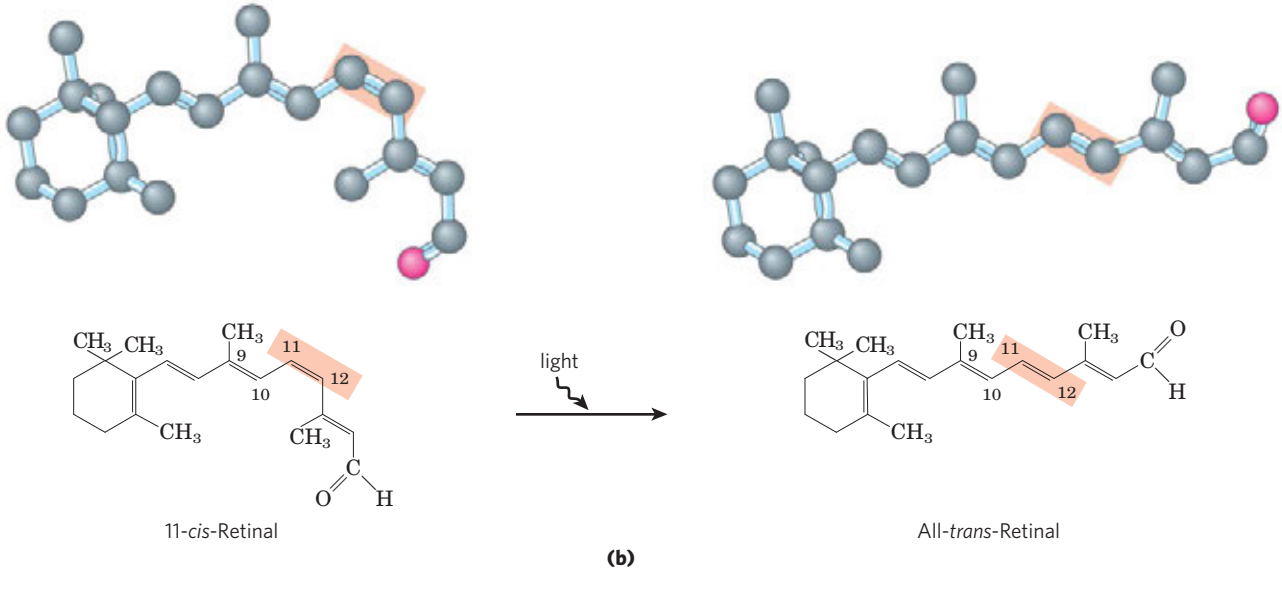


FIGURE 1–19 Configurations of geometric isomers. (a) Isomers such as maleic acid (maleate at pH 7) and fumaric acid (fumarate) cannot be interconverted without breaking covalent bonds, which requires the input of much more energy than the average kinetic energy of molecules at physiological temperatures. **(b)** In the vertebrate retina, the initial event in light detection is the absorption of visible light by 11-*cis*-retinal. The energy of the absorbed light (about 250 kJ/mol) converts 11-*cis*-retinal to all-*trans*-retinal, triggering electrical changes in the retinal cell that lead to a nerve impulse. (Note that the hydrogen atoms are omitted from the ball-and-stick models of the retinals.)

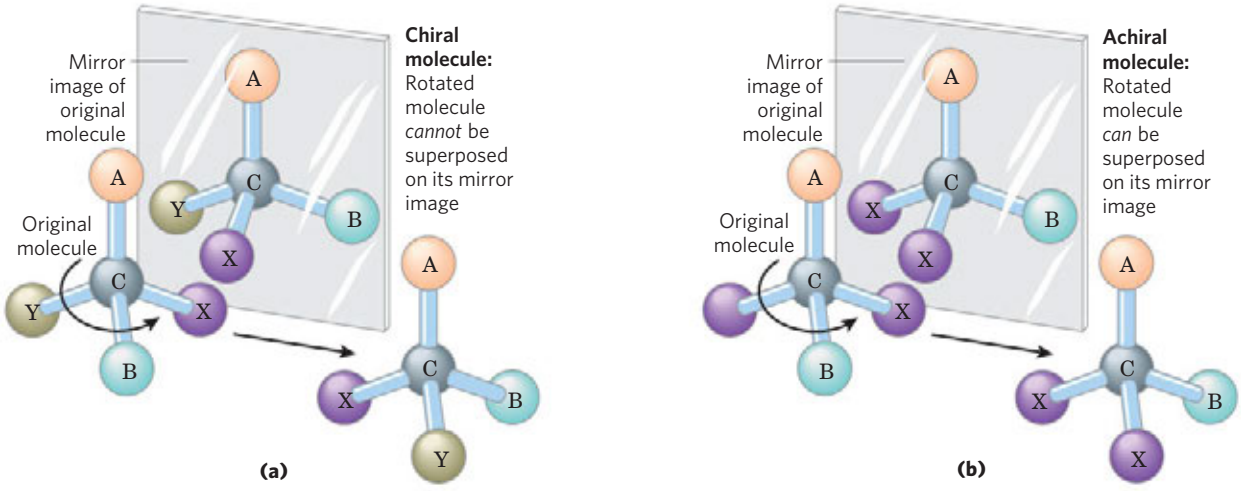


FIGURE 1–20 Molecular asymmetry: chiral and achiral molecules. (a) When a carbon atom has four different substituent groups (A, B, X, Y), they can be arranged in two ways that represent nonsuperposable mirror images of each other (enantiomers). This asymmetric carbon atom is called a chiral atom or chiral center. **(b)** When a tetrahedral carbon has only three dissimilar groups (i.e., the same group occurs twice), only one configuration is possible and the molecule is symmetric, or achiral. In this case the molecule is superposable on its mirror image: the molecule on the left can be rotated counterclockwise (when looking down the vertical bond from A to C) to create the molecule in the mirror.

isomers, or **cis-trans isomers**; they differ in the arrangement of their substituent groups with respect to the nonrotating double bond (Latin *cis*, "on this side"—groups on the same side of the double bond; *trans*, "across"—groups on opposite sides). Maleic acid (maleate at the neutral pH of cytoplasm) is the cis isomer and fumaric acid (fumarate) the trans isomer; each is a well-defined compound that can be separated from the other, and each has its own unique chemical properties. A binding site (on an enzyme, for example) that is complementary to one of these molecules would not be complementary to the other, which explains why the two compounds have distinct biological roles despite their similar chemical makeup.

In the second type of stereoisomer, four different substituents bonded to a tetrahedral carbon atom may be arranged in two different ways in space—that is, have two configurations **(Fig. 1–20)**—yielding two stereoisomers that have similar or identical chemical properties but differ in certain physical and biological properties. A carbon atom with four different substituents is said to be asymmetric, and asymmetric carbons are called **chiral centers** (Greek *chiros*, "hand"; some stereoisomers are related structurally as the right hand is to the left). A molecule with only one chiral carbon can have two stereoisomers; when two or more (n) chiral carbons are present, there can be 2^n stereoisomers. Stereoisomers that are mirror images of each other are called **enantiomers** (Fig. 1–20). Pairs of stereoisomers that are not mirror images of each other are called **diastereomers (Fig. 1–21)**.

As Louis Pasteur first observed in 1843 (Box 1–2), enantiomers have nearly identical chemical reactivities but differ in a characteristic physical property: their interaction with plane-polarized light. In separate solutions, two enantiomers rotate the plane of plane-polarized light in opposite directions, but an equimolar solution of the two enantiomers (a **racemic mixture**) shows no optical rotation. Compounds without chiral centers do not rotate the plane of plane-polarized light.

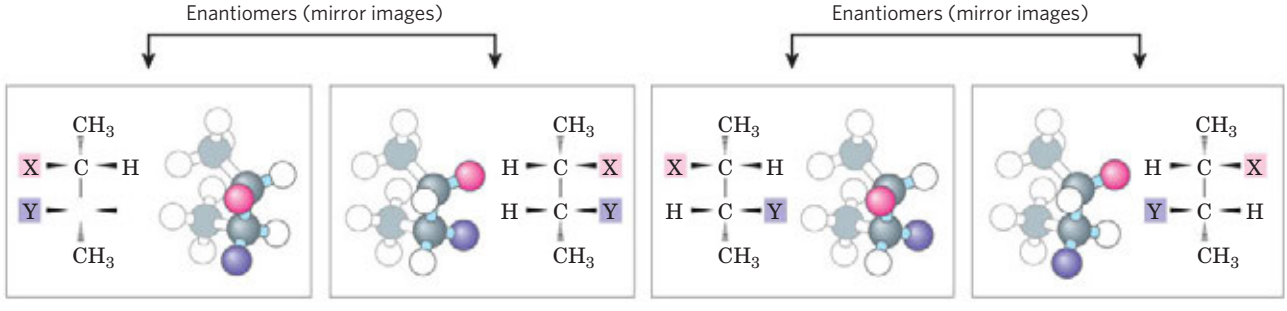


FIGURE 1–21 Enantiomers and diastereomers. There are four different stereoisomers of 2,3-disubstituted butane (n = 2 asymmetric carbons, hence 2^n = 4 stereoisomers). Each is shown in a box as a perspective formula and a ball-and-stick model, which has been rotated to allow the reader to view all the groups. Two pairs of stereoisomers are mirror images of each other, or enantiomers. All other possible pairs are not mirror images and so are diastereomers.

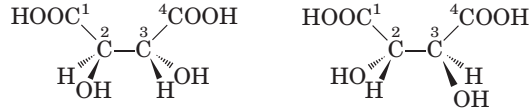
BOX 1–2 Louis Pasteur and Optical Activity: *In Vino, Veritas*

Louis Pasteur encountered the phenomenon of **optical activity** in 1843, during his investigation of the crystalline sediment that accumulated in wine casks (a form of tartaric acid called paratartaric acid—also called racemic acid, from Latin *racemus*, "bunch of grapes"). He used fine forceps to separate two types of crystals identical in shape but mirror images of each other. Both types proved to have all the chemical properties of tartaric acid, but in solution one type rotated plane-polarized light to the left (levorotatory), the other to the right (dextrorotatory). Pasteur later described the experiment and its interpretation:

Louis Pasteur, 1822–1895

> In isomeric bodies, the elements and the proportions in which they are combined are the same, only the arrangement of the atoms is different . . . We know, on the one hand, that the molecular arrangements of the two tartaric acids are asymmetric, and, on the other hand, that these arrangements are absolutely identical, excepting that they exhibit asymmetry in opposite directions. Are the atoms of the dextro acid grouped in the form of a right-handed spiral, or are they placed at the apex of an irregular tetrahedron, or are they disposed according to this or that asymmetric arrangement? We do not know.*

Now we do know. X-ray crystallographic studies in 1951 confirmed that the levorotatory and dextrorotatory forms of tartaric acid are mirror images of each other at the molecular level and established the absolute configuration of each (Fig. 1). The same approach has been used to demonstrate that although the amino acid alanine has two stereoisomeric forms (designated D and L), alanine in proteins exists exclusively in one form (the L isomer; see Chapter 3).



(2*R*,3*R*)-Tartaric acid (dextrorotatory) | (2*S*,3*S*)-Tartaric acid (levorotatory)

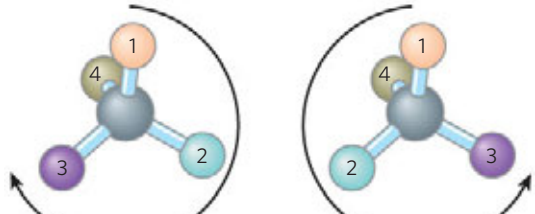
FIGURE 1 Pasteur separated crystals of two stereoisomers of tartaric acid and showed that solutions of the separated forms rotated plane-polarized light to the same extent but in opposite directions. These dextrorotatory and levorotatory forms were later shown to be the (*R*,*R*) and (*S*,*S*) isomers represented here. The RS system of nomenclature is explained in the text.

*From Pasteur's lecture to the Société Chimique de Paris in 1883, quoted in DuBos, R. (1976) *Louis Pasteur: Free Lance of Science*, p. 95, Charles Scribner's Sons, New York.

KEY CONVENTION: Given the importance of stereochemistry in reactions between biomolecules (see below), biochemists must name and represent the structure of each biomolecule so that its stereochemistry is unambiguous. For compounds with more than one chiral center, the most useful system of nomenclature is the RS system. In this system, each group attached to a chiral carbon is assigned a *priority*. The priorities of some common substituents are

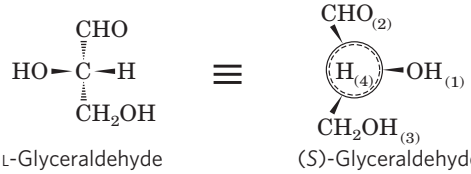
$$-OCH_3 > -OH > -NH_2 > -COOH > -CHO > -CH_2OH > -CH_3 > -H$$

For naming in the RS system, the chiral atom is viewed with the group of lowest priority (4 in the following diagram) pointing away from the viewer. If the priority of the other three groups (1 to 3) decreases in clockwise order, the configuration is (*R*) (Latin *rectus*, "right"); if counterclockwise, the configuration is (*S*) (Latin *sinister*, "left"). In this way each chiral carbon is designated either (*R*) or (*S*), and the inclusion of these designations in the name of the compound provides an unambiguous description of the stereochemistry at each chiral center.



Clockwise (*R*) | Counterclockwise (*S*)

Another naming system for stereoisomers, the D and L system, is described in Chapter 3. A molecule with a single chiral center (the two isomers of glyceraldehyde, for example) can be named unambiguously by either system.



L-Glyceraldehyde | (*S*)-Glyceraldehyde ■

Distinct from configuration is molecular **conformation**, the spatial arrangement of substituent groups that, without breaking any bonds, are free to assume different positions in space because of the freedom of rotation about single bonds. In the simple hydrocarbon ethane, for example, there is nearly complete freedom of rotation around the C—C bond. Many different, interconvertible conformations of ethane are possible, depending on the degree of rotation **(Fig. 1–22)**. Two conformations are of special interest: the staggered, which is more stable than all others and thus predominates, and the eclipsed, which is least stable. We cannot isolate either of these conformational forms, because they are freely interconvertible. However, when one or more of the hydrogen atoms on each carbon is replaced by a functional group that is either very large or electrically charged, freedom of rotation around the C—C bond is hindered. This limits the number of stable conformations of the ethane derivative.

Interactions between Biomolecules Are Stereospecific

When biomolecules interact, the "fit" between them must be stereochemically correct. The three-dimensional structure of biomolecules large and small—the combination of configuration and conformation—is of the utmost importance in their biological interactions: reactant with its enzyme, hormone with its receptor on a cell surface, antigen with its specific antibody, for example **(Fig. 1–23)**. The study of biomolecular stereochemistry, with precise physical methods, is an important part of modern research on cell structure and biochemical function.

In living organisms, chiral molecules are usually present in only one of their chiral forms. For example, the amino acids in proteins occur only as their L isomers; glucose occurs only as its D isomer. (The conventions for naming stereoisomers of the amino acids are described in Chapter 3; those for sugars, in Chapter 7. The RS system, described above, is the most useful for some biomolecules.) In contrast, when a compound with an asymmetric carbon atom is chemically synthesized in the laboratory, the reaction usually produces all possible chiral forms: a mixture of the D and L forms, for example. Living cells produce only one chiral form of a biomolecule because the enzymes that synthesize that molecule are also chiral.

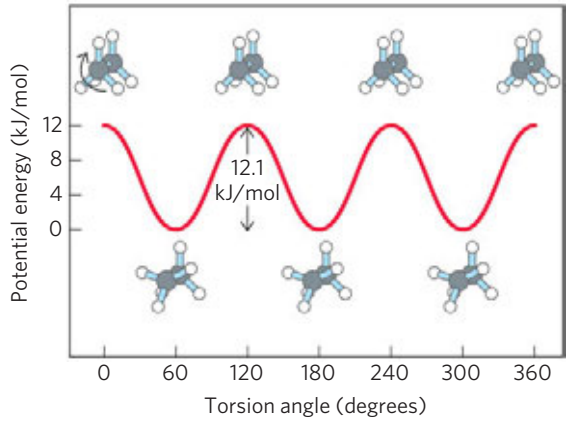


FIGURE 1–22 Conformations. Many conformations of ethane are possible because of freedom of rotation around the C—C bond. In the ball-and-stick model, when the front carbon atom (as viewed by the reader) with its three attached hydrogens is rotated relative to the rear carbon atom, the potential energy of the molecule rises to a maximum in the fully eclipsed conformation (torsion angle 0°, 120°, etc.), then falls to a minimum in the fully staggered conformation (torsion angle 60°, 180°, etc.). Because the energy differences are small enough to allow rapid interconversion of the two forms (millions of times per second), the eclipsed and staggered forms cannot be separately isolated.

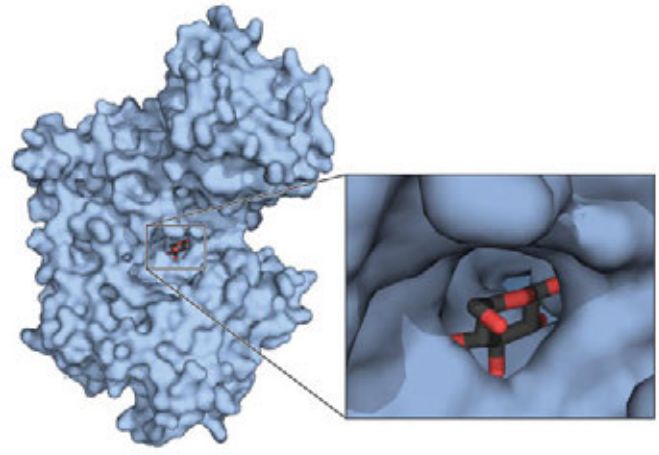


FIGURE 1–23 Complementary fit between a macromolecule and a small molecule. A glucose molecule fits into a pocket on the surface of the enzyme hexokinase (PDB ID 3B8A), and is held in this orientation by several noncovalent interactions between the protein and the sugar. This representation of the hexokinase molecule is produced with software that can calculate the shape of the outer surface of a macromolecule, defined either by the van der Waals radii of all the atoms in the molecule or by the "solvent exclusion volume," the volume a water molecule cannot penetrate.

Stereospecificity, the ability to distinguish between stereoisomers, is a property of enzymes and other proteins and a characteristic feature of the molecular logic of living cells. If the binding site on a protein is complementary to one isomer of a chiral compound, it will not be complementary to the other isomer, for the same reason that a left glove does not fit a right hand. Some striking examples of the ability of biological systems to distinguish stereoisomers are shown in **Figure 1–24**.

The common classes of chemical reactions encountered in biochemistry are described in Chapter 13, as an introduction to the reactions of metabolism.

SUMMARY 1.2 Chemical Foundations

- Because of its bonding versatility, carbon can produce a broad array of carbon–carbon skeletons with a variety of functional groups; these groups give biomolecules their biological and chemical personalities.
- A nearly universal set of about a thousand small molecules is found in living cells; the interconversions of these molecules in the central

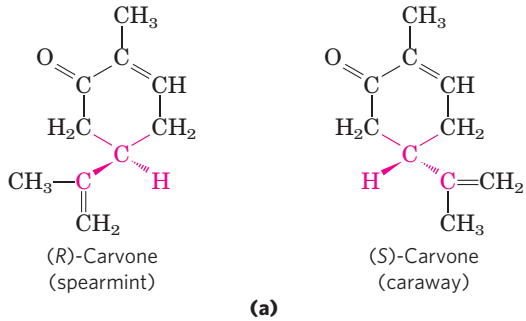


FIGURE 1–24 Stereoisomers have different effects in humans. (a) Two stereoisomers of carvone: (*R*)-carvone (isolated from spearmint oil) has the characteristic fragrance of spearmint; (*S*)-carvone (from caraway seed oil) smells like caraway. **(b)** Aspartame, the artificial sweetener sold under the trade name NutraSweet, is easily distinguishable by taste receptors from its bitter-tasting stereoisomer, although the two differ only in the configuration at one of the two chiral carbon atoms. **(c)** The antidepressant medication citalopram (trade name Celexa), a selective serotonin reuptake inhibitor, is a racemic mixture of these two steroisomers, but only (*S*)-citalopram has the therapeutic effect. A stereochemically pure preparation of (*S*)-citalopram (escitalopram oxalate) is sold under the trade name Lexapro. As you might predict, the effective dose of Lexapro is one-half the effective dose of Celexa.

metabolic pathways have been conserved in evolution.

- Proteins and nucleic acids are linear polymers of simple monomeric subunits; their sequences contain the information that gives each molecule its three-dimensional structure and its biological functions.
- Molecular configuration can be changed only by breaking covalent bonds. For a carbon atom with four different substituents (a chiral carbon), the substituent groups can be arranged in two different ways, generating stereoisomers with distinct properties. Only one stereoisomer is biologically active. Molecular conformation is the position of atoms in space that can be changed by rotation about single bonds, without breaking covalent bonds.
- Interactions between biological molecules are almost invariably stereospecific: they require a close fit between complementary structures in the interacting molecules.

1.3 Physical Foundations

Living cells and organisms must perform work to stay alive and to reproduce themselves. The synthetic reactions that occur within cells, like the synthetic processes in any factory, require the input of energy. Energy input is also needed in the motion of a bacterium or an Olympic sprinter, in the flashing of a firefly or the electrical discharge of an eel. And the storage and expression of information require energy, without which structures rich in information inevitably become disordered and meaningless.

In the course of evolution, cells have developed highly efficient mechanisms for coupling the energy obtained from sunlight or chemical fuels to the many energy-requiring processes they must carry out. One goal of biochemistry is to understand, in quantitative and chemical terms, the means by which energy is extracted, stored, and channeled into useful work in living cells. We can consider cellular energy conversions—like all other energy conversions—in the context of the laws of thermodynamics.

Living Organisms Exist in a Dynamic Steady State, Never at Equilibrium with Their Surroundings

The molecules and ions contained within a living organism differ in kind and in concentration from those in the organism's surroundings. A paramecium in a pond, a shark in the ocean, a bacterium in the soil, an apple tree in an orchard—all are different in composition from their surroundings and, once they have reached maturity, maintain a more or less constant composition in the face of constantly changing surroundings.

Although the characteristic composition of an organism changes little through time, the population of molecules within the organism is far from static. Small molecules, macromolecules, and supramolecular complexes are continuously synthesized and broken down in chemical reactions that involve a constant flux of mass and energy through the system. The hemoglobin molecules carrying oxygen from your lungs to your brain at this moment were synthesized within the past month; by next month they will have been degraded and entirely replaced by new hemoglobin molecules. The glucose you ingested with your most recent meal is now circulating in your bloodstream; before the day is over these particular glucose molecules will have been converted into something else—carbon dioxide or fat, perhaps—and will have been replaced with a fresh supply of glucose, so that your blood glucose concentration is more or less constant over the whole day. The amounts of hemoglobin and glucose in the blood remain nearly constant because the rate of synthesis or intake of each just balances the rate of its breakdown, consumption, or conversion into some other product. The constancy of concentration is the result of a *dynamic steady state*, a steady state that is far from equilibrium. Maintaining this steady state requires the constant investment of energy; when a cell can no longer obtain energy, it dies and begins to decay toward equilibrium with its surroundings. We consider below exactly what is meant by "steady state" and "equilibrium."

Organisms Transform Energy and Matter from Their Surroundings

For chemical reactions occurring in solution, we can define a **system** as all the constituent reactants and products, the solvent that contains them, and the immediate atmosphere—in short, everything within a defined region of space. The system and its surroundings together constitute the **universe**. If the system exchanges neither matter nor energy with its surroundings, it is said to be **isolated**. If the system exchanges energy but not matter with its surroundings, it is a **closed** system; if it exchanges both energy and matter with its surroundings, it is an **open** system.

A living organism is an open system; it exchanges both matter and energy with its surroundings. Organisms obtain energy from their surroundings in two ways: (1) they take up chemical fuels (such as glucose) from the environment and extract energy by oxidizing them (see Box 1–3, Case 2); or (2) they absorb energy from sunlight.

The first law of thermodynamics describes the principle of the conservation of energy: *in any physical or chemical change, the total amount of energy in the universe remains constant, although the form of the energy may change*. This means that while energy is "used" by a system, it is not "used up"; rather it is converted from one form into another—from potential energy in chemical bonds, say, into kinetic energy of heat and motion. Cells are consummate transducers of energy, capable of interconverting chemical, electromagnetic, mechanical, and osmotic energy with great efficiency **(Fig. 1–25)**.

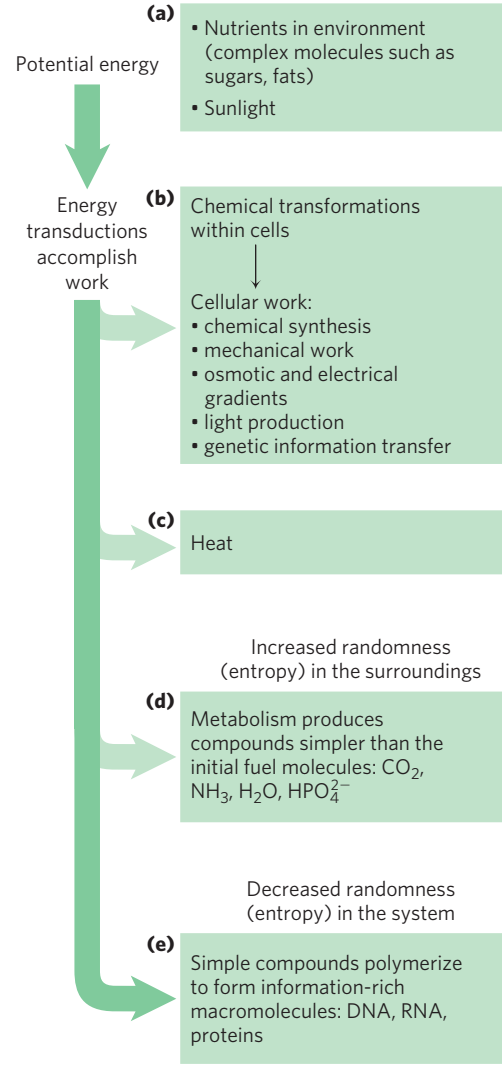


FIGURE 1–25 Some energy transformations in living organisms. As metabolic energy is spent to do cellular work, the randomness of the system plus surroundings (expressed quantitatively as entropy) increases as the potential energy of complex nutrient molecules decreases. **(a)** Living organisms extract energy from their surroundings; **(b)** convert some of it into useful forms of energy to produce work; **(c)** return some energy to the surroundings as heat; and **(d)** release end-product molecules that are less well organized than the starting fuel, increasing the entropy of the universe. One effect of all these transformations is **(e)** increased order (decreased randomness) in the system in the form of complex macromolecules. We return to a quantitative treatment of entropy in Chapter 13.

BOX 1–3 Entropy: Things Fall Apart

The term "entropy," which literally means "a change within," was first used in 1851 by Rudolf Clausius, one of the formulators of the second law of thermodynamics. It refers to the randomness or disorder of the components of a chemical system. Entropy is a central concept in biochemistry; life requires continual maintenance of order in the face of nature's tendency to increase randomness. A rigorous quantitative definition of entropy involves statistical and probability considerations. However, its nature can be illustrated qualitatively by three simple examples, each demonstrating one aspect of entropy. The key descriptors of entropy are *randomness* and *disorder*, manifested in different ways.

Case 1: The Teakettle and the Randomization of Heat

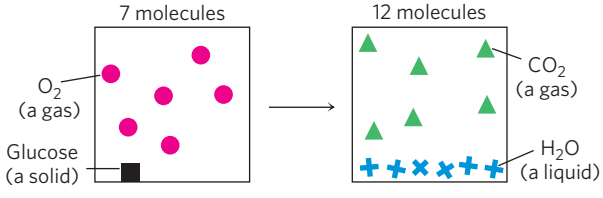
We know that steam generated from boiling water can do useful work. But suppose we turn off the burner under a teakettle full of water at 100 °C (the "system") in the kitchen (the "surroundings") and allow the teakettle to cool. As it cools, no work is done, but heat passes from the teakettle to the surroundings, raising the temperature of the surroundings (the kitchen) by an infinitesimally small amount until complete equilibrium is attained. At this point all parts of the teakettle and the kitchen are at precisely the same temperature. The free energy that was once concentrated in the teakettle of hot water at 100 °C, *potentially* capable of doing work, has disappeared. Its equivalent in heat energy is still present in the teakettle + kitchen (i.e., the "universe") but has become completely randomized throughout. This energy is no longer available to do work because there is no temperature differential within the kitchen. Moreover, the increase in entropy of the kitchen (the surroundings) is irreversible. We know from everyday experience that heat never spontaneously passes back from the kitchen into the teakettle to raise the temperature of the water to 100 °C again.

Case 2: The Oxidation of Glucose

Entropy is a state not only of energy but of matter. Aerobic (heterotrophic) organisms extract free energy from glucose obtained from their surroundings by oxidizing the glucose with O_2, also obtained from the surroundings. The end products of this oxidative metabolism, CO_2 and H_2O, are returned to the surroundings. In this process the surroundings undergo an increase in entropy, whereas the organism itself remains in a steady state and undergoes no change in its internal order. Although some entropy arises from the dissipation of heat, entropy also arises from another kind of disorder, illustrated by the equation for the oxidation of glucose:

$$C_6H_{12}O_6 + 6O_2 \longrightarrow 6CO_2 + 6H_2O$$

We can represent this schematically as



The Flow of Electrons Provides Energy for Organisms

Nearly all living organisms derive their energy, directly or indirectly, from the radiant energy of sunlight. In the photoautotrophs, light-driven splitting of water during photosynthesis releases its electrons for the reduction of CO_2 and the release of O_2 into the atmosphere:

$$6CO_2 + 6H_2O \xrightarrow{\text{light}} C_6H_{12}O_6 + 6O_2$$

(light-driven reduction of CO_2)

Nonphotosynthetic organisms (chemotrophs) obtain the energy they need by oxidizing the energy-rich products of photosynthesis stored in plants, then passing the electrons thus acquired to atmospheric O_2 to form water, CO_2, and other end products, which are recycled in the environment:

$$C_6H_{12}O_6 + 6O_2 \longrightarrow 6CO_2 + 6H_2O + \text{energy}$$

(energy-yielding oxidation of glucose)

Thus autotrophs and heterotrophs participate in global cycles of O_2 and CO_2, driven ultimately by sunlight, making these two large groups of organisms interdependent. Virtually all energy transductions in cells can be traced to this flow of electrons from one molecule to another, in a "downhill" flow from higher to lower electrochemical potential; as such, this is formally analogous to the flow of electrons in a battery-driven electric circuit. All these reactions involved in electron flow are **oxidation-reduction reactions**: one reactant is oxidized (loses electrons) as another is reduced (gains electrons).

Creating and Maintaining Order Requires Work and Energy

As we've noted, DNA, RNA, and proteins are informational macromolecules; the precise sequence of their monomeric subunits contains information, just as the letters in this sentence do. In addition to using chemical energy to form the covalent bonds between these subunits, the cell must invest energy to order the subunits in their correct sequence. It is extremely improbable that amino acids in a mixture would spontaneously condense into a single type of protein, with a unique sequence. This would represent increased order in a population of molecules; but

The atoms contained in 1 molecule of glucose plus 6 molecules of oxygen, a total of 7 molecules, are more randomly dispersed by the oxidation reaction and are now present in a total of 12 molecules ($6CO_2 + 6H_2O$).

Whenever a chemical reaction results in an increase in the number of molecules—or when a solid substance is converted into liquid or gaseous products, which allow more freedom of molecular movement than solids—molecular disorder, and thus entropy, increases.

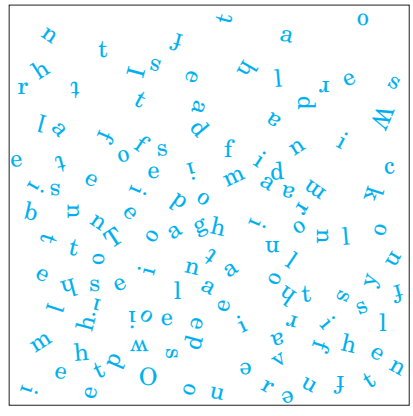
Case 3: Information and Entropy

The following short passage from *Julius Caesar*, Act IV, Scene 3, is spoken by Brutus, when he realizes that he must face Mark Antony's army. It is an information-rich nonrandom arrangement of 125 letters of the English alphabet:

> There is a tide in the affairs of men,
> Which, taken at the flood, leads on to fortune;
> Omitted, all the voyage of their life
> Is bound in shallows and in miseries.

In addition to what this passage says overtly, it has many hidden meanings. It not only reflects a complex sequence of events in the play, it also echoes the play's ideas on conflict, ambition, and the demands of leadership. Permeated with Shakespeare's understanding of human nature, it is very rich in information.

However, if the 125 letters making up this quotation were allowed to fall into a completely random, chaotic pattern, as shown in the following box, they would have no meaning whatsoever.



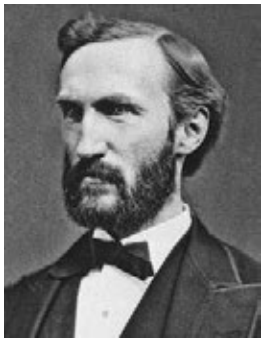
In this form the 125 letters contain little or no information, but they are very rich in entropy. Such considerations have led to the conclusion that information is a form of energy; information has been called "negative entropy." In fact, the branch of mathematics called information theory, which is basic to the programming logic of computers, is closely related to thermodynamic theory. Living organisms are highly ordered, nonrandom structures, immensely rich in information and thus entropy-poor.

according to the second law of thermodynamics, the tendency in nature is toward ever-greater disorder in the universe: *the total entropy of the universe is continually increasing*. To bring about the synthesis of macromolecules from their monomeric units, free energy must be supplied to the system (in this case, the cell).

KEY CONVENTION: The randomness or disorder of the components of a chemical system is expressed as **entropy**, ***S*** (Box 1–3). Any change in randomness of the system is expressed as entropy change, ΔS, which by convention has a positive value when randomness increases. J. Willard Gibbs, who developed the theory of energy changes during chemical reactions, showed that the **free-energy content**, ***G***, of any closed system can be defined in terms of three quantities: **enthalpy**, ***H***, reflecting the number and kinds of bonds; entropy, S; and the absolute temperature, T (in Kelvin). The definition of free energy is $G = H - TS$. When a chemical reaction occurs at constant temperature, the **free-energy change**, ΔG, is determined by the enthalpy change, ΔH, reflecting the kinds and numbers of chemical bonds and noncovalent interactions broken and formed, and the entropy change, ΔS, describing the change in the system's randomness:

$$\Delta G = \Delta H - T\Delta S$$

where, by definition, ΔH is negative for a reaction that releases heat, and ΔS is positive for a reaction that increases the system's randomness. ■



J. Willard Gibbs, 1839–1903

A process tends to occur spontaneously only if ΔG is negative (if free energy is *released* in the process). Yet cell function depends largely on molecules, such as proteins and nucleic acids, for which the free energy of formation is positive: the molecules are less stable and more highly ordered than a mixture of their monomeric components. To carry out these thermodynamically unfavorable, energy-requiring (**endergonic**) reactions, cells couple them to other reactions that liberate free

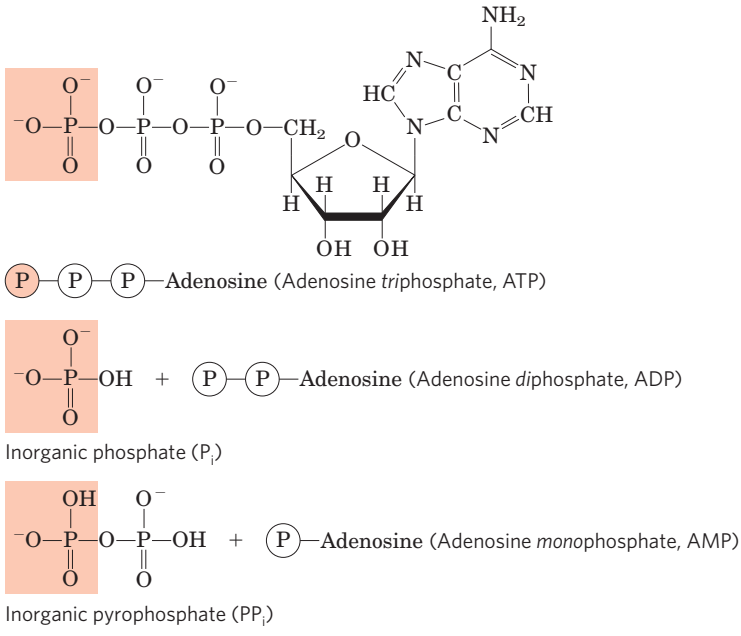


FIGURE 1–26 Adenosine triphosphate (ATP) provides energy. Here, each Ⓟ represents a phosphoryl group. The removal of the terminal phosphoryl group (shaded light red) of ATP, by breakage of a phosphoanhydride bond to generate adenosine diphosphate (ADP) and inorganic phosphate ion (HPO_4^{2-}), is highly exergonic, and this reaction is coupled to many endergonic reactions in the cell (as in the example in Fig. 1–27b). ATP also provides energy for many cellular processes by undergoing cleavage that releases the two terminal phosphates as inorganic pyrophosphate ($H_2P_2O_7^{2-}$), often abbreviated PP_i.

energy (**exergonic** reactions), so that the overall process is exergonic: the *sum* of the free-energy changes is negative.

The usual source of free energy in coupled biological reactions is the energy released by breakage of phosphoanhydride bonds such as those in adenosine triphosphate (ATP; **Fig. 1–26**) and guanosine triphosphate (GTP). Here, each Ⓟ represents a phosphoryl group:

Amino acids → protein	ΔG_1 is positive (endergonic)
ATP → AMP + Ⓟ—Ⓟ	ΔG_2 is negative (exergonic)
[or ATP → ADP + Ⓟ]	

When these reactions are coupled, the sum of ΔG_1 and ΔG_2 is negative—the overall process is exergonic. By this coupling strategy, cells are able to synthesize and maintain the information-rich polymers essential to life.

Energy Coupling Links Reactions in Biology

The central issue in *bioenergetics* (the study of energy transformations in living systems) is the means by which energy from fuel metabolism or light capture is coupled to a cell's energy-requiring reactions. In thinking about energy coupling, it is useful to consider a simple mechanical example, as shown in **Figure 1–27a**. An object at the top of an inclined plane has a certain amount of potential energy as a result of its elevation. It tends to slide down the plane, losing its potential energy of position as it approaches the ground. When an appropriate string-and-pulley device couples the falling object to another, smaller object, the spontaneous downward

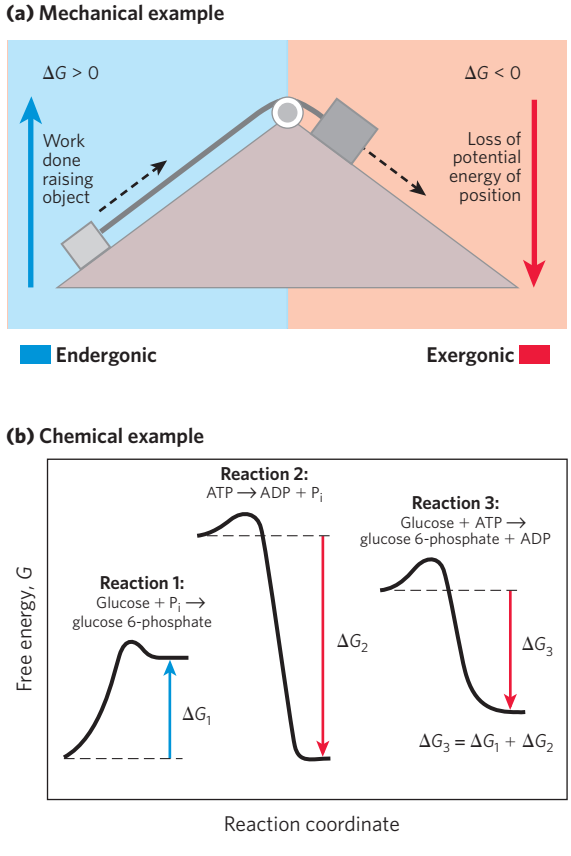


FIGURE 1–27 Energy coupling in mechanical and chemical processes. **(a)** The downward motion of an object releases potential energy that can do mechanical work. The potential energy made available by spontaneous downward motion, an exergonic process (red), can be coupled to the endergonic upward movement of another object (blue). **(b)** In reaction 1, the formation of glucose 6-phosphate from glucose and inorganic phosphate (P_i) yields a product of higher energy than the two reactants. For this endergonic reaction, ΔG is positive. In reaction 2, the exergonic breakdown of adenosine triphosphate (ATP) has a large, negative free-energy change (ΔG_2). The third reaction is the sum of reactions 1 and 2, and the free-energy change, ΔG_3, is the arithmetic sum of ΔG_1 and ΔG_2. Because ΔG_3 is negative, the overall reaction is exergonic and proceeds spontaneously.

motion of the larger can lift the smaller, accomplishing a certain amount of work. The amount of energy available to do work is the **free-energy change**, ΔG; this is always somewhat less than the theoretical amount of energy released, because some energy is dissipated as the heat of friction. The greater the elevation of the larger object, the greater the energy released (ΔG) as the object slides downward and the greater the amount of work that can be accomplished. The larger object can lift the smaller only because, at the outset, the larger object was *far from its equilibrium position:* it had at some earlier point been elevated above the ground, in a process that itself required the input of energy.

How does this apply in chemical reactions? In closed systems, chemical reactions proceed spontaneously until **equilibrium** is reached. When a system is at equilibrium, the rate of product formation exactly equals the rate at which product is converted to reactant. Thus there is no net change in the concentration of reactants and products. The energy change as the system moves from its initial state to equilibrium, with no changes in temperature or pressure, is given by the free-energy change, ΔG. The magnitude of ΔG depends on the particular chemical reaction *and on how far from equilibrium the system is initially*. Each compound involved in a chemical reaction contains a certain amount of potential energy, related to the kind and number of its bonds. In reactions that occur spontaneously, the products have less free energy than the reactants, thus the reaction releases free energy, which is then available to do work. Such reactions are exergonic; the decline in free energy from reactants to products is expressed as a negative value. Endergonic reactions require an input of energy, and their ΔG values are positive. As in mechanical processes, only part of the energy released in exergonic chemical reactions can be used to accomplish work. In living systems some energy is dissipated as heat or lost to increasing entropy.

K_{eq} and $\Delta G°$ Are Measures of a Reaction's Tendency to Proceed Spontaneously

The tendency of a chemical reaction to go to completion can be expressed as an equilibrium constant. For the reaction in which a moles of A react with b moles of B to give c moles of C and d moles of D,

$$aA + bB \longrightarrow cC + dD$$

the equilibrium constant, K_{eq}, is given by

$$K_{eq} = \frac{[C]_{eq}^{c}[D]_{eq}^{d}}{[A]_{eq}^{a}[B]_{eq}^{b}}$$

where $[A]_{eq}$ is the concentration of A, $[B]_{eq}$ the concentration of B, and so on, *when the system has reached equilibrium*. A large value of K_{eq} means the reaction tends to proceed until the reactants are almost completely converted into the products.

WORKED EXAMPLE 1–1 Are ATP and ADP at Equilibrium in Cells?

The equilibrium constant, K_{eq}, for the following reaction is 2×10^5 M:

$$ATP \longrightarrow ADP + HPO_4^{2-}$$

If the measured cellular concentrations are [ATP] = 5 mM, [ADP] = 0.5 mM, and $[P_i]$ = 5 mM, is this reaction at equilibrium in living cells?

Solution: The definition of the equilibrium constant for this reaction is:

$$K_{eq} = [ADP][P_i]/[ATP]$$

From the measured cellular concentrations given above, we can calculate the mass-action ratio, Q:

$$Q = [ADP][P_i]/[ATP] = [0.5 \text{ mM}][5 \text{ mM}]/[5 \text{ mM}]$$
$$= 0.5 \text{ mM} = 5 \times 10^{-4} \text{ M}$$

This value is *far* from the equilibrium constant for the reaction (2×10^5 M), so the reaction is *very far* from equilibrium in cells. [ATP] is far higher, and [ADP] is far lower, than is expected at equilibrium. How can a cell hold its [ATP]/[ADP] ratio so far from equilibrium? It does so by continuously extracting energy (from nutrients such as glucose) and using it to make ATP from ADP and P_i.

WORKED EXAMPLE 1–2 Is the Hexokinase Reaction at Equilibrium in Cells?

For the reaction catalyzed by the enzyme hexokinase:

$$\text{Glucose} + ATP \longrightarrow \text{glucose 6-phosphate} + ADP$$

the equilibrium constant, K_{eq}, is 7.8×10^2. In living *E. coli* cells, [ATP] = 5 mM; [ADP] = 0.5 mM; [glucose] = 2 mM; and [glucose 6-phosphate] = 1 mM. Is the reaction at equilibrium in *E. coli*?

Solution: At equilibrium,

$$K_{eq} = 7.8 \times 10^2 = [ADP][\text{glucose 6-phosphate}]/[ATP][\text{glucose}]$$

In living cells, [ADP][glucose 6-phosphate]/[ATP][glucose]= [0.5 mM][1 mM]/[5 mM][2 mM] = 0.05. The reaction is therefore *far* from equilibrium: the cellular concentrations of products (glucose 6-phosphate and ADP) are much lower than expected at equilibrium, and those of the reactants are much higher. The reaction therefore tends strongly to go to the right.

Gibbs showed that ΔG (the actual free-energy change) for any chemical reaction is a function of the **standard free-energy change**, $\Delta G°$—a constant that is characteristic of each specific reaction—and a term

that expresses the initial concentrations of reactants and products:

$$\Delta G = \Delta G^\circ + RT \ln \frac{[C]_i^c [D]_i^d}{[A]_i^a [B]_i^b} \qquad (1\text{–}1)$$

where $[A]_i$ is the initial concentration of A, and so forth; R is the gas constant; and T is the absolute temperature.

ΔG is a measure of the distance of a system from its equilibrium position. When a reaction has reached equilibrium, no driving force remains and it can do no work: $\Delta G = 0$. For this special case, $[A]_i = [A]_{eq}$, and so on, for all reactants and products, and

$$\frac{[C]_i^c [D]_i^d}{[A]_i^a [B]_i^b} = \frac{[C]_{eq}^c [D]_{eq}^d}{[A]_{eq}^a [B]_{eq}^b}$$

Substituting 0 for ΔG and K_{eq} for $[C]_i^c[D]_i^d/[A]_i^a[B]_i^b$ in Equation 1–1, we obtain the relationship

$$\Delta G^\circ = -RT \ln K_{eq}$$

from which we see that ΔG° is simply a second way (besides K_{eq}) of expressing the driving force on a reaction. Because K_{eq} is experimentally measurable, we have a way of determining ΔG°, the thermodynamic constant characteristic of each reaction.

The units of ΔG° and ΔG are joules per mole (or calories per mole). When $K_{eq} >> 1$, ΔG° is large and negative; when $K_{eq} << 1$, ΔG° is large and positive. From a table of experimentally determined values of either K_{eq} or ΔG°, we can see at a glance which reactions tend to go to completion and which do not.

One caution about the interpretation of ΔG°: *thermodynamic* constants such as this show where the final equilibrium for a reaction lies but tell us nothing about how fast that equilibrium will be achieved. The rates of reactions are governed by the parameters of *kinetics*, a topic we consider in detail in Chapter 6.

In biological organisms, just as in the mechanical example in Figure 1–27a, an exergonic reaction can be coupled to an endergonic reaction to drive otherwise unfavorable reactions. Figure 1–27b (a type of graph called a reaction coordinate diagram) illustrates this principle for the conversion of glucose to glucose 6-phosphate, the first step in the pathway for oxidation of glucose. The simplest way to produce glucose 6-phosphate would be:

Reaction 1: Glucose + P_i ⟶ glucose 6-phosphate
(endergonic; ΔG_1 is positive)

(Here, P_i is an abbreviation for inorganic phosphate, HPO_4^{2-}. Don't be concerned about the structure of these compounds now; we describe them in detail later in the book.) This reaction does not occur spontaneously; ΔG_1 is positive. A second, very exergonic reaction can occur in all cells:

Reaction 2: ATP ⟶ ADP + P_i
(exergonic; ΔG_2 is negative)

These two chemical reactions share a common intermediate, P_i, which is consumed in reaction 1 and produced in reaction 2. The two reactions can therefore be coupled in the form of a third reaction, which we can write as the sum of reactions 1 and 2, with the common intermediate, P_i, omitted from both sides of the equation:

Reaction 3: Glucose + ATP ⟶
glucose 6-phosphate + ADP

Because more energy is released in reaction 2 than is consumed in reaction 1, the free-energy change for reaction 3, ΔG_3, is negative, and the synthesis of glucose 6-phosphate can therefore occur by reaction 3.

WORKED EXAMPLE 1–3 Standard Free-Energy Changes Are Additive

Given that the standard free-energy change for the reaction glucose + P_i ⟶ glucose 6-phosphate is 13.8 kJ/mol, and the standard free-energy change for the reaction ATP ⟶ ADP + P_i is −30.5 kJ/mol, what is the free-energy change for the reaction glucose + ATP ⟶ glucose 6-phosphate + ADP?

Solution: We can write the equation for this reaction as the sum of two other reactions:

(1) Glucose + P_i ⟶ glucose 6-phosphate $\Delta G_1^\circ = 13.8$ kJ/mol
(2) ATP ⟶ ADP + P_i $\Delta G_2^\circ = -30.5$ kJ/mol

Sum: Glucose + ATP ⟶ glucose 6-phosphate + ADP
$\Delta G_{Sum}^\circ = -16.7$ kJ/mol

The standard free-energy change for two reactions that sum to a third is simply the sum of the two individual reactions. A negative value for ΔG° (–16.7 kJ/mol) indicates that the reaction will tend to occur spontaneously.

The coupling of exergonic and endergonic reactions through a shared intermediate is central to the energy exchanges in living systems. As we shall see, reactions that break down ATP (such as reaction 2 in Fig. 1–27b) release energy that drives many endergonic processes in cells. ATP breakdown in cells is exergonic because *all living cells maintain a concentration of ATP far above its equilibrium concentration*. It is this disequilibrium that allows ATP to serve as the major carrier of chemical energy in all cells. As we shall see in more detail in Chapter 13, it is not the mere breakdown of ATP that provides energy to drive endergonic reactions; rather, it is the *transfer of a phosphoryl group* from ATP to another small molecule (glucose in the case above) that conserves some of the chemical potential originally in ATP.

WORKED EXAMPLE 1–4 Energetic Cost of ATP Synthesis

If the equilibrium constant, K_{eq}, for the reaction

$$ATP \longrightarrow ADP + P_i$$

is 2.22×10^5 M, calculate the standard free-energy change, $\Delta G°$, for the *synthesis* of ATP from ADP and P_i at 25 °C.

Solution: First calculate $\Delta G°$ for the reaction above.

$$\begin{aligned}\Delta G° &= -RT \ln K_{eq} \\ &= -(8.315 \text{ J/mol} \cdot \text{K})(298 \text{ K})(\ln 2.22 \times 10^5) \\ &= -30.5 \text{ kJ/mol}\end{aligned}$$

This is the standard free-energy change for the *breakdown* of ATP to ADP and P_i. The standard free-energy change for the *reverse* of a reaction has the same absolute value, but the opposite sign. The standard free-energy change for the reverse of the above reaction is therefore 30.5 kJ/mol. So, to synthesize 1 mol of ATP under standard conditions (25 °C, 1 M concentrations of ATP, ADP, P_i), at least 30.5 kJ of energy must be supplied. The actual free-energy change in cells—approximately 50 kJ/mol—is greater than this because the concentrations of ATP, ADP, and P_i in cells are not the standard 1 M (see Worked Example 13–2, p. 519).

WORKED EXAMPLE 1–5 Standard Free-Energy Change for Synthesis of Glucose 6-Phosphate

What is the standard free-energy change, $\Delta G°$, under physiological conditions (*E. coli* grows in the human gut, at 37 °C) for the following reaction?

$$\text{Glucose} + ATP \longrightarrow \text{glucose 6-phosphate} + ADP$$

Solution: We have the relationship $\Delta G° = -RT \ln K_{eq}$ and the value of K_{eq} for this reaction, 7.8×10^2. Substituting the values of R, T, and K_{eq} into this equation gives:

$$\Delta G° = -(8.315 \text{ J/mol} \cdot \text{K})(310 \text{ K})(\ln 7.8 \times 10^2) = -17 \text{ kJ/mol}$$

Notice that this value is slightly different from that in Worked Example 1–3. In that calculation we assumed a temperature of 25 °C (298 K), whereas in this calculation we used the physiological temperature of 37 °C (310 K).

Enzymes Promote Sequences of Chemical Reactions

All biological macromolecules are much less thermodynamically stable than their monomeric subunits, yet they are *kinetically stable*: their *uncatalyzed* breakdown occurs so slowly (over years rather than seconds) that, on a time scale that matters for the organism, these molecules are stable. Virtually every chemical reaction in a cell occurs at a significant rate only because of the presence

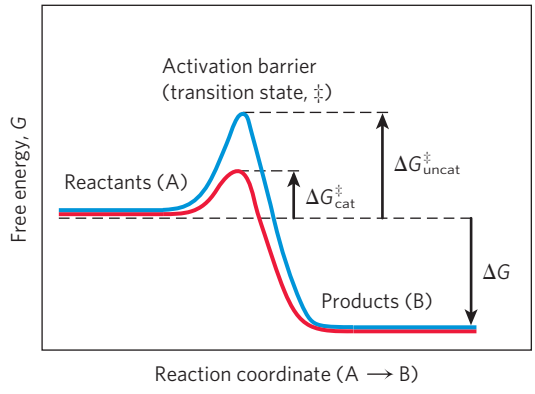


FIGURE 1–28 Energy changes during a chemical reaction. An activation barrier, representing the transition state (see Chapter 6), must be overcome in the conversion of reactants (A) into products (B), even though the products are more stable than the reactants, as indicated by a large, negative free-energy change (ΔG). The energy required to overcome the activation barrier is the activation energy ($\Delta G^{\ddagger}$). Enzymes catalyze reactions by lowering the activation barrier. They bind the transition-state intermediates tightly, and the binding energy of this interaction effectively reduces the activation energy from $\Delta G^{\ddagger}_{uncat}$ (blue curve) to $\Delta G^{\ddagger}_{cat}$ (red curve). (Note that activation energy is *not* related to free-energy change, ΔG.)

of **enzymes**—biocatalysts that, like all other catalysts, greatly enhance the rate of specific chemical reactions without being consumed in the process.

The path from reactant(s) to product(s) almost invariably involves an energy barrier, called the activation barrier **(Fig. 1–28)**, that must be surmounted for any reaction to proceed. The breaking of existing bonds and formation of new ones generally requires, first, a distortion of the existing bonds to create a **transition state** of higher free energy than either reactant or product. The highest point in the reaction coordinate diagram represents the transition state, and the difference in energy between the reactant in its ground state and in its transition state is the **activation energy**, $\Delta G^{\ddagger}$. An enzyme catalyzes a reaction by providing a more comfortable fit for the transition state: a surface that complements the transition state in stereochemistry, polarity, and charge. The binding of enzyme to the transition state is exergonic, and the energy released by this binding reduces the activation energy for the reaction and greatly increases the reaction rate.

A further contribution to catalysis occurs when two or more reactants bind to the enzyme's surface close to each other and with stereospecific orientations that favor the reaction. This increases by orders of magnitude the probability of productive collisions between reactants. As a result of these factors and several others, discussed in Chapter 6, enzyme-catalyzed reactions commonly proceed at rates greater than 10^{12} times faster than the uncatalyzed reactions. (That is a *million million* times faster!)

Cellular catalysts are, with some notable exceptions, proteins. (Some RNA molecules have enzymatic activity,

as discussed in Chapters 26 and 27.) Again with a few exceptions, each enzyme catalyzes a specific reaction, and each reaction in a cell is catalyzed by a different enzyme. Thousands of different enzymes are therefore required by each cell. The multiplicity of enzymes, their specificity (the ability to discriminate between reactants), and their susceptibility to regulation give cells the capacity to lower activation barriers selectively. This selectivity is crucial for the effective regulation of cellular processes. By allowing specific reactions to proceed at significant rates at particular times, enzymes determine how matter and energy are channeled into cellular activities.

The thousands of enzyme-catalyzed chemical reactions in cells are functionally organized into many sequences of consecutive reactions, called **pathways**, in which the product of one reaction becomes the reactant in the next. Some pathways degrade organic nutrients into simple end products in order to extract chemical energy and convert it into a form useful to the cell; together these degradative, free-energy-yielding reactions are designated **catabolism**. The energy released by catabolic reactions drives the synthesis of ATP. As a result, the cellular concentration of ATP is far above its equilibrium concentration, so that ΔG for ATP breakdown is large and negative. Similarly, catabolism results in the production of the reduced electron carriers NADH and NADPH, both of which can donate electrons in processes that generate ATP or drive reductive steps in biosynthetic pathways.

Other pathways start with small precursor molecules and convert them to progressively larger and more complex molecules, including proteins and nucleic acids. Such synthetic pathways, which invariably require the input of energy, are collectively designated **anabolism**. The overall network of enzyme-catalyzed pathways, both catabolic and anabolic, constitutes cellular **metabolism**. ATP (and the energetically equivalent nucleoside triphosphates cytidine triphosphate (CTP), uridine triphosphate (UTP), and guanosine triphosphate (GTP)) is the connecting link between the catabolic and anabolic components of this network (shown schematically in **Fig. 1–29**). The pathways of enzyme-catalyzed reactions that act on the main constituents of cells—proteins, fats, sugars, and nucleic acids—are virtually identical in all living organisms.

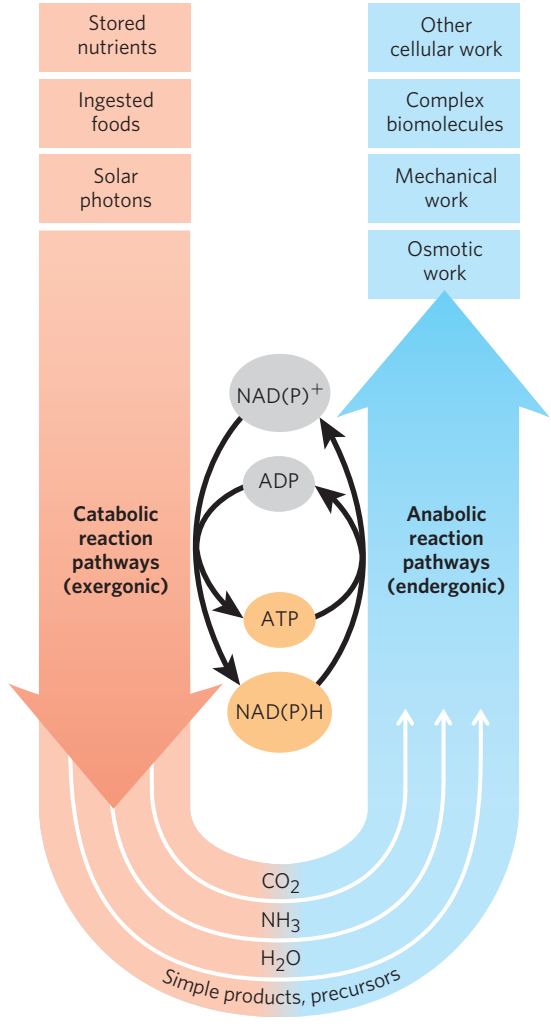
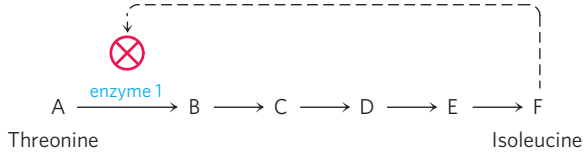


FIGURE 1–29 The central roles of ATP and NAD(P)H in metabolism. ATP is the shared chemical intermediate linking energy-releasing and energy-consuming cellular processes. Its role in the cell is analogous to that of money in an economy: it is "earned/produced" in exergonic reactions and "spent/consumed" in endergonic ones. NAD(P)H (nicotinamide adenine dinucleotide (phosphate)) is an electron-carrying cofactor that collects electrons from oxidative reactions and then donates them in a wide variety of reduction reactions in biosynthesis. Present in relatively low concentrations, these cofactors essential to anabolic reactions must be constantly regenerated by catabolic reactions.

Metabolism Is Regulated to Achieve Balance and Economy

Not only do living cells simultaneously synthesize thousands of different kinds of carbohydrate, fat, protein, and nucleic acid molecules and their simpler subunits, but they do so in the precise proportions required by the cell under any given circumstance. For example, during rapid cell growth the precursors of proteins and nucleic acids must be made in large quantities, whereas in nongrowing cells the requirement for these precursors is much reduced. Key enzymes in each metabolic pathway are regulated so that each type of precursor molecule is produced in a quantity appropriate to the current requirements of the cell.

Consider the pathway in *E. coli* that leads to the synthesis of the amino acid isoleucine, a constituent of proteins. The pathway has five steps catalyzed by five different enzymes (A through F represent the intermediates in the pathway):



If a cell begins to produce more isoleucine than it needs for protein synthesis, the unused isoleucine accumulates and the increased concentration inhibits the catalytic activity of the first enzyme in the pathway, immediately slowing the production of isoleucine. Such **feedback inhibition** keeps the production and utilization of each metabolic intermediate in balance. (Throughout the book, we will use ⊗ to indicate inhibition of an enzymatic reaction.)

Although the concept of discrete pathways is an important tool for organizing our understanding of metabolism, it is an oversimplification. There are thousands of metabolic intermediates in a cell, many of which are part of more than one pathway. Metabolism would be better represented as a web of interconnected and interdependent pathways. A change in the concentration of any one metabolite would start a ripple effect, influencing the flow of materials through other pathways. The task of understanding these complex interactions among intermediates and pathways in quantitative terms is daunting, but the new emphasis on **systems biology**, discussed in Chapter 15, has begun to offer important insights into the overall regulation of metabolism.

Cells also regulate the synthesis of their own catalysts, the enzymes, in response to increased or diminished need for a metabolic product; this is the substance of Chapter 28. The expression of genes (the translation from information in DNA to active protein in the cell) and synthesis of enzymes are other layers of metabolic control in the cell. All layers must be taken into account when describing the overall control of cellular metabolism.

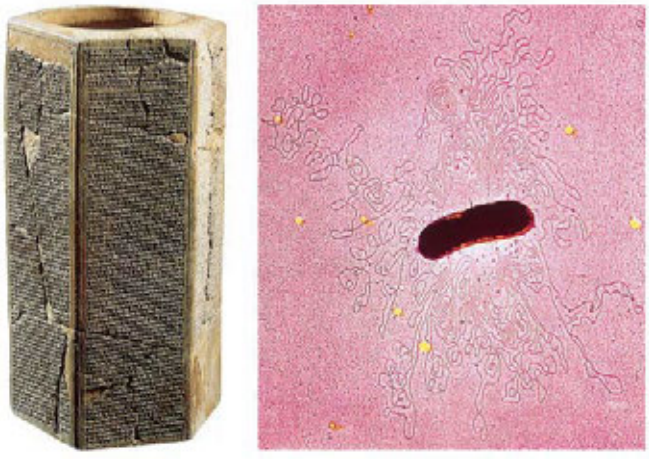
SUMMARY 1.3 Physical Foundations

- Living cells are open systems, exchanging matter and energy with their surroundings, extracting and channeling energy to maintain themselves in a dynamic steady state distant from equilibrium. Energy is obtained from sunlight or chemical fuels by converting the energy from electron flow into the chemical bonds of ATP.
- The tendency for a chemical reaction to proceed toward equilibrium can be expressed as the free-energy change, ΔG, which has two components: enthalpy change, ΔH, and entropy change, ΔS. These variables are related by the equation $\Delta G = \Delta H - T\Delta S$.
- When ΔG of a reaction is negative, the reaction is exergonic and tends to go toward completion; when ΔG is positive, the reaction is endergonic and tends to go in the reverse direction. When two reactions can be summed to yield a third reaction, the ΔG for this overall reaction is the sum of the ΔGs of the two separate reactions.
- The reactions converting ATP to P_i and ADP or to AMP and PP_i are highly exergonic (large negative ΔG). Many endergonic cellular reactions are driven by coupling them, through a common intermediate, to these highly exergonic reactions.
- The standard free-energy change for a reaction, $\Delta G°$, is a physical constant that is related to the equilibrium constant by the equation $\Delta G° = -RT \ln K_{eq}$.
- Most cellular reactions proceed at useful rates only because enzymes are present to catalyze them. Enzymes act in part by stabilizing the transition state, reducing the activation energy, $\Delta G^{\ddagger}$, and increasing the reaction rate by many orders of magnitude. The catalytic activity of enzymes in cells is regulated.
- Metabolism is the sum of many interconnected reaction sequences that interconvert cellular metabolites. Each sequence is regulated to provide what the cell needs at a given time and to expend energy only when necessary.

1.4 Genetic Foundations

Perhaps the most remarkable property of living cells and organisms is their ability to reproduce themselves for countless generations with nearly perfect fidelity. This continuity of inherited traits implies constancy, over millions of years, in the structure of the molecules that contain the genetic information. Very few historical records of civilization, even those etched in copper or carved in stone **(Fig. 1–30)**, have survived for a thousand years. But there is good evidence that the genetic instructions in living organisms have remained nearly unchanged over very much longer periods; many bacteria have nearly the same size, shape, and internal structure and contain the same kinds of precursor molecules and enzymes as bacteria that lived nearly four billion years ago. This continuity of structure and composition is the result of continuity in the structure of the genetic material.

Among the seminal discoveries in biology in the twentieth century were the chemical nature and the three-dimensional structure of the genetic material, **deoxyribonucleic acid**, **DNA**. The sequence of the monomeric subunits, the nucleotides (strictly, deoxyribonucleotides, as discussed below), in this linear polymer encodes the instructions for forming all other cellular components and provides a template for the production of identical DNA molecules to be distributed to progeny when a cell divides. The perpetuation of a biological species requires that its genetic information be maintained in a stable form, expressed accurately in the form of gene products, and reproduced with a minimum of errors. Effective storage, expression, and reproduction of the genetic message define individual species, distinguish them from one another, and assure their continuity over successive generations.



(a) (b)

FIGURE 1–30 Two ancient scripts. (a) The Prism of Sennacherib, inscribed in about 700 BCE, describes in characters of the Assyrian language some historical events during the reign of King Sennacherib. The Prism contains about 20,000 characters, weighs about 50 kg, and has survived almost intact for about 2,700 years. **(b)** The single DNA molecule of the bacterium *E. coli*, leaking out of a disrupted cell, is hundreds of times longer than the cell itself and contains all the encoded information necessary to specify the cell's structure and functions. The bacterial DNA contains about 4.6 million characters (nucleotides), weighs less than 10^{-10} g, and has undergone only relatively minor changes during the past several million years. (The yellow spots and dark specks in this colorized electron micrograph are artifacts of the preparation.)

Genetic Continuity Is Vested in Single DNA Molecules

DNA is a long, thin, organic polymer, the rare molecule that is constructed on the atomic scale in one dimension (width) and the human scale in another (length: a molecule of DNA can be many centimeters long). A human sperm or egg, carrying the accumulated hereditary information of billions of years of evolution, transmits this inheritance in the form of DNA molecules, in which the linear sequence of covalently linked nucleotide subunits encodes the genetic message.

Usually when we describe the properties of a chemical species, we describe the average behavior of a very large number of identical molecules. While it is difficult to predict the behavior of any single molecule in a collection of, say, a picomole (about 6×10^{11} molecules) of a compound, the *average* behavior of the molecules is predictable because so many molecules enter into the average. Cellular DNA is a remarkable exception. The DNA that is the entire genetic material of *E. coli* is a *single molecule* containing 4.64 million nucleotide pairs. That single molecule must be replicated perfectly in every detail if an *E. coli* cell is to give rise to identical progeny by cell division; there is no room for averaging in this process! The same is true for all cells. A human sperm brings to the egg that it fertilizes just one molecule of DNA in each of its 23 different chromosomes, to combine with just one DNA molecule in each corresponding chromosome in the egg. The result of this union is very highly predictable: an embryo with all of its ~25,000 genes, constructed of 3 billion nucleotide pairs, intact. An amazing chemical feat.

WORKED EXAMPLE 1–6 Fidelity of DNA Replication

Calculate the number of times the DNA of a modern *E. coli* cell has been copied accurately since its earliest bacterial precursor cell arose about 3.5 billion years ago. Assume for simplicity that over this time period *E. coli* has undergone, on average, one cell division every 12 hours (this is an overestimate for modern bacteria, but probably an underestimate for ancient bacteria).

Solution:

$$(1 \text{ generation}/12 \text{ hr})(24 \text{ hr/d})(365 \text{ d/yr})(3.5 \times 10^9 \text{ yr}) = 2.6 \times 10^{12} \text{ generations.}$$

A single page of this book contains about 5,000 characters, so the entire book contains about 5 million characters. The chromosome of *E. coli* also contains about 5 million characters (nucleotide pairs). If you made a handwritten copy of this book and then passed on the copy to a classmate to copy by hand, and this copy were then copied by a third classmate, and so on, how closely would each successive copy of the book resemble the original? Now, imagine the textbook that would result from hand-copying this one a few trillion times!

The Structure of DNA Allows for Its Replication and Repair with Near-Perfect Fidelity

The capacity of living cells to preserve their genetic material and to duplicate it for the next generation results from the structural complementarity between the two strands of the DNA molecule **(Fig. 1–31)**. The basic unit of DNA is a linear polymer of four different monomeric subunits, **deoxyribonucleotides**, arranged in a precise linear sequence. It is this linear sequence that encodes the genetic information. Two of these polymeric strands are twisted about each other to form the DNA double helix, in which each deoxyribonucleotide in one strand pairs specifically with a complementary deoxyribonucleotide in the opposite strand. Before a cell divides, the two DNA strands separate and each serves as a template for the synthesis of a new, complementary strand, generating two identical double-helical molecules, one for each daughter cell. If either strand is damaged at any time, continuity of information is assured by the information present in the other strand, which can act as a template for repair of the damage.

The Linear Sequence in DNA Encodes Proteins with Three-Dimensional Structures

The information in DNA is encoded in its linear (one-dimensional) sequence of deoxyribonucleotide subunits,

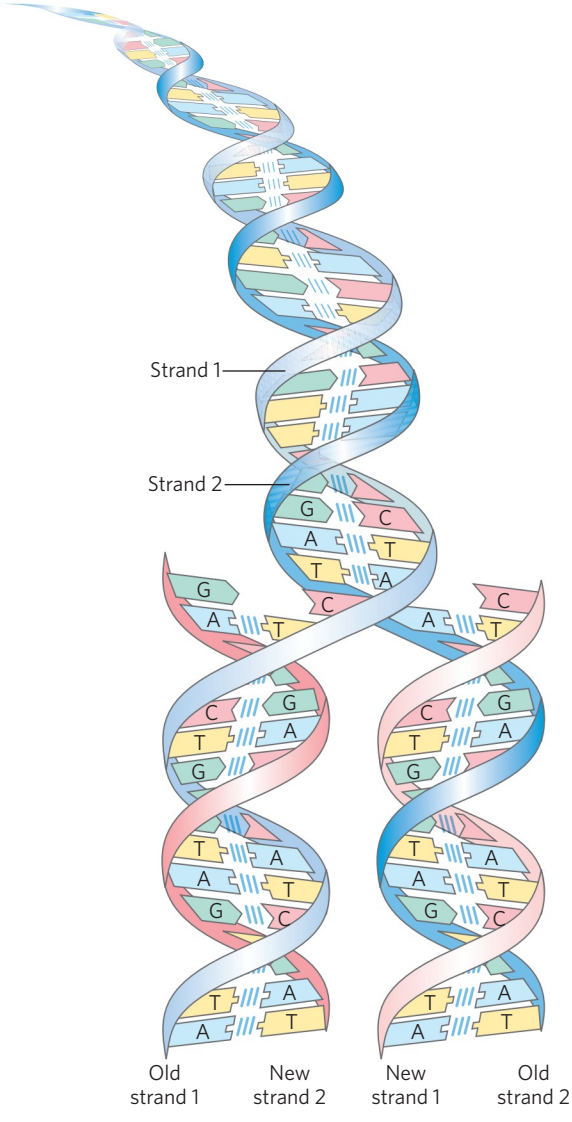


FIGURE 1–31 Complementarity between the two strands of DNA. DNA is a linear polymer of covalently joined deoxyribonucleotides of four types: deoxyadenylate (A), deoxyguanylate (G), deoxycytidylate (C), and deoxythymidylate (T). Each nucleotide, with its unique three-dimensional structure, can associate very specifically but noncovalently with one other nucleotide in the complementary chain: A always associates with T, and G with C. Thus, in the double-stranded DNA molecule, the entire sequence of nucleotides in one strand is *complementary* to the sequence in the other. The two strands, held together by hydrogen bonds (represented here by vertical light blue lines) between each pair of complementary nucleotides, twist about each other to form the DNA double helix. In DNA replication, the two strands (blue) separate and two new strands (pink) are synthesized, each with a sequence complementary to one of the original strands. The result is two double-helical molecules, each identical to the original DNA.

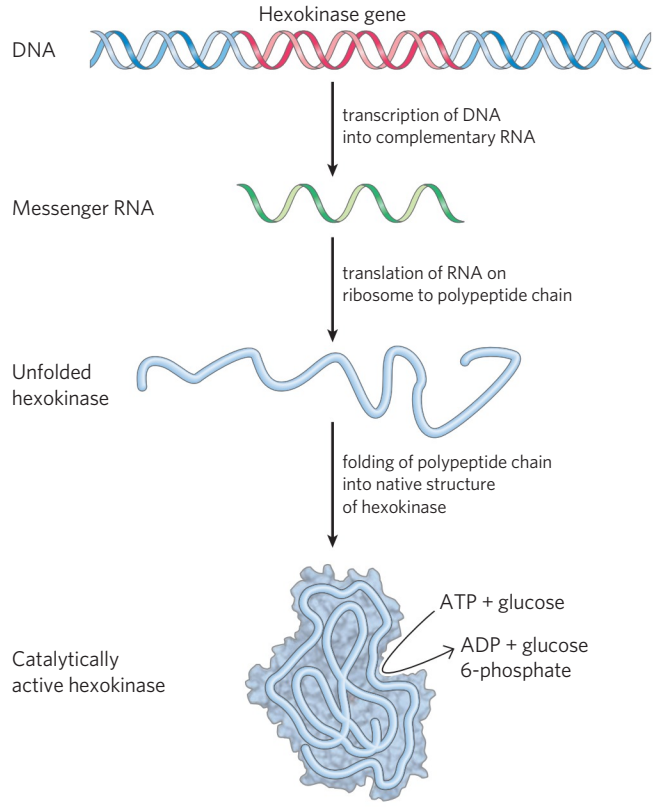


FIGURE 1–32 DNA to RNA to protein to enzyme (hexokinase). The linear sequence of deoxyribonucleotides in the DNA (the gene) that encodes the protein hexokinase is first transcribed into a ribonucleic acid (RNA) molecule with the complementary ribonucleotide sequence. The RNA sequence (messenger RNA) is then translated into the linear protein chain of hexokinase, which folds into its native three-dimensional shape, most likely aided by molecular chaperones. Once in its native form, hexokinase acquires its catalytic activity: it can catalyze the phosphorylation of glucose, using ATP as the phosphoryl group donor.

but the expression of this information results in a three-dimensional cell. This change from one to three dimensions occurs in two phases. A linear sequence of deoxyribonucleotides in DNA codes (through an intermediary, RNA) for the production of a protein with a corresponding linear sequence of amino acids **(Fig. 1–32)**. The protein folds into a particular three-dimensional shape, determined by its amino acid sequence and stabilized primarily by noncovalent interactions. Although the final shape of the folded protein is dictated by its amino acid sequence, the folding is aided by "molecular chaperones" (see Fig. 4–30). The precise three-dimensional structure, or **native conformation**, of the protein is crucial to its function.

Once in its native conformation, a protein may associate noncovalently with other macromolecules (other proteins, nucleic acids, or lipids) to form supramolecular complexes such as chromosomes, ribosomes, and membranes. The individual molecules of these complexes have specific, high-affinity binding sites for each other, and within the cell they spontaneously self-assemble into functional complexes.

Although the amino acid sequences of proteins carry all necessary information for achieving the proteins' native conformation, accurate folding and self-assembly also require the right cellular environment—pH, ionic strength, metal ion concentrations, and so forth. Thus DNA sequence alone is not enough to form and maintain a fully functioning cell.

SUMMARY 1.4 Genetic Foundations

- Genetic information is encoded in the linear sequence of four types of deoxyribonucleotides in DNA.
- The double-helical DNA molecule contains an internal template for its own replication and repair.
- DNA molecules are extraordinarily large, with molecular weights in the millions or billions.
- Despite the enormous size of DNA, the sequence of nucleotides in it is very precise, and the maintenance of this precise sequence over very long times is the basis for genetic continuity in organisms.
- The linear sequence of amino acids in a protein, which is encoded in the DNA of the gene for that protein, produces a protein's unique three-dimensional structure—a process also dependent on environmental conditions.
- Individual macromolecules with specific affinity for other macromolecules self-assemble into supramolecular complexes.

1.5 Evolutionary Foundations

Nothing in biology makes sense except in the light of evolution.

—*Theodosius Dobzhansky,* The American Biology Teacher, *March 1973*

Great progress in biochemistry and molecular biology in recent decades has amply confirmed the validity of Dobzhansky's striking generalization. The remarkable similarity of metabolic pathways and gene sequences across the three domains of life argues strongly that all modern organisms are derived from a common evolutionary progenitor by a series of small changes (mutations), each of which conferred a selective advantage to some organism in some ecological niche.

Changes in the Hereditary Instructions Allow Evolution

Despite the near-perfect fidelity of genetic replication, infrequent unrepaired mistakes in the DNA replication process lead to changes in the nucleotide sequence of DNA, producing a genetic **mutation** and changing the instructions for a cellular component. Incorrectly repaired damage to one of the DNA strands has the same effect. Mutations in the DNA handed down to offspring—that is, mutations carried in the reproductive cells—may be harmful or even lethal to the new organism or cell; they may, for example, cause the synthesis of a defective enzyme that is not able to catalyze an essential metabolic reaction. Occasionally, however, a mutation *better* equips an organism or cell to survive in its environment **(Fig. 1–33)**. The mutant enzyme might have acquired a slightly different specificity, for example, so that it is now able to use some compound that the cell was previously

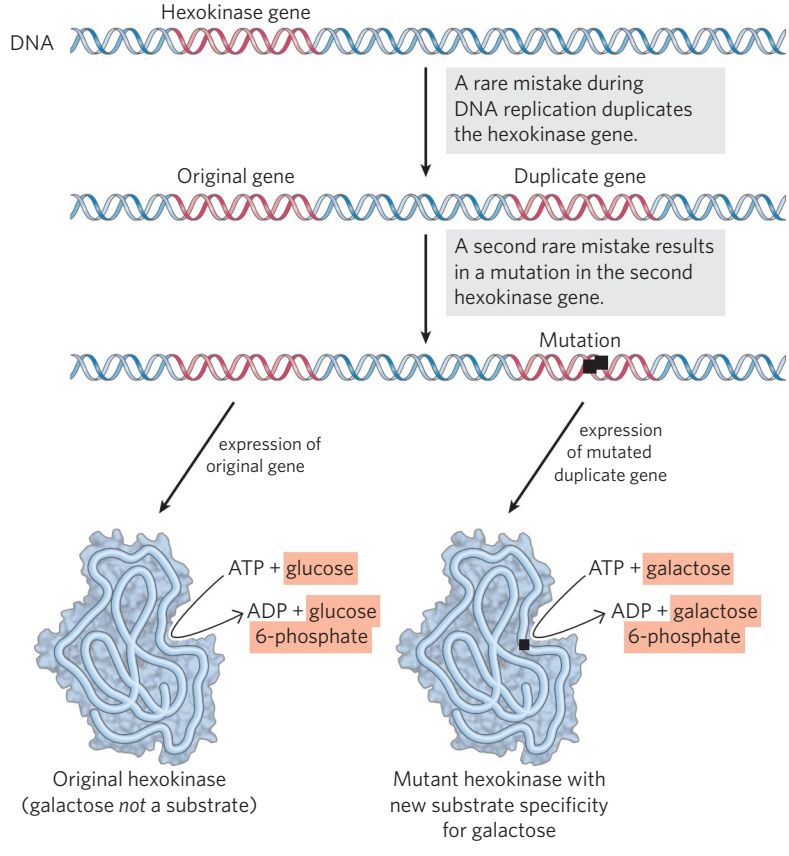


FIGURE 1–33 Gene duplication and mutation: one path to generate new enzymatic activities. In this example, the single hexokinase gene in a hypothetical organism might occasionally, by accident, be copied twice during DNA replication, such that the organism has two full copies of the gene, one of which is superfluous. Over many generations, as the DNA with two hexokinase genes is repeatedly duplicated, rare mistakes occur, leading to changes in the nucleotide sequence of the superfluous gene and thus of the protein that it encodes. In a few very rare cases, the altered protein produced from this mutant gene can bind a new substrate—galactose in our hypothetical case. The cell containing the mutant gene has acquired a new capability (metabolism of galactose), which may allow it to survive in an ecological niche that provides galactose but not glucose. If no gene duplication precedes mutation, the original function of the gene product is lost.

unable to metabolize. If a population of cells were to find itself in an environment where that compound was the only or the most abundant available source of fuel, the mutant cell would have a selective advantage over the other, unmutated (**wild-type**) cells in the population. The mutant cell and its progeny would survive and prosper in the new environment, whereas wild-type cells would starve and be eliminated. This is what Darwin meant by natural selection—what is sometimes summarized as "survival of the fittest."

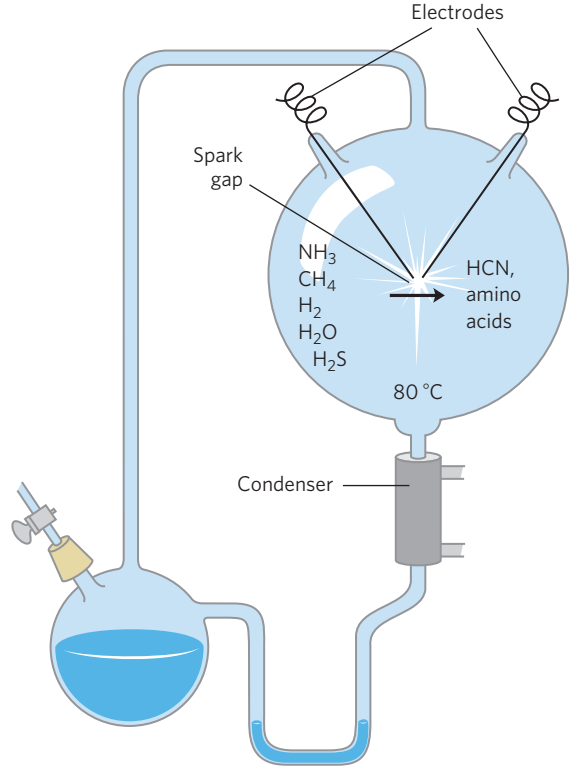
Occasionally, a second copy of a whole gene is introduced into the chromosome as a result of defective replication of the chromosome. The second copy is superfluous, and mutations in this gene will not be deleterious; it becomes a means by which the cell may evolve, by producing a new gene with a new function while retaining the original gene and gene function. Seen in this light, the DNA molecules of modern organisms are historical documents, records of the long journey from the earliest cells to modern organisms. The historical accounts in DNA are not complete, however; in the course of evolution, many mutations must have been erased or written over. But DNA molecules are the best source of biological history that we have. The frequency of errors in DNA replication represents a balance between too many errors, which would yield nonviable daughter cells, and too few, which would prevent the genetic variation that allows survival of mutant cells in new ecological niches.

Several billion years of natural selection have refined cellular systems to take maximum advantage of the chemical and physical properties of available raw materials. Chance genetic mutations occurring in individuals in a population, combined with natural selection, have resulted in the evolution of the enormous variety of species we see today, each adapted to its particular ecological niche.

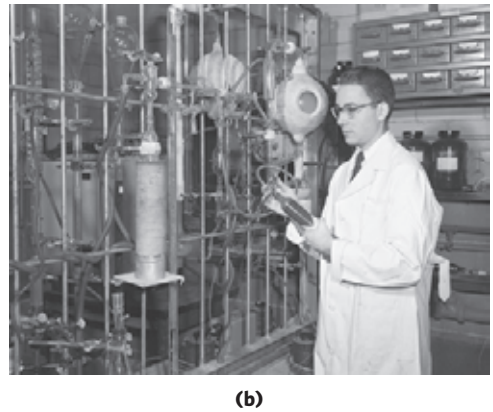
Biomolecules First Arose by Chemical Evolution

In our account thus far we have passed over the first chapter of the story of evolution: the appearance of the first living cell. Apart from their occurrence in living organisms, organic compounds, including the basic biomolecules such as amino acids and carbohydrates, are found in only trace amounts in the Earth's crust, the sea, and the atmosphere. How did the first living organisms acquire their characteristic organic building blocks? According to one hypothesis, these compounds were created by the effects of powerful environmental forces—ultraviolet irradiation, lightning, or volcanic eruptions—on the gases in the prebiotic Earth's atmosphere, and on inorganic solutes in superheated thermal vents deep in the ocean.

This hypothesis was tested in a classic experiment on the abiotic (nonbiological) origin of organic biomolecules carried out in 1953 by Stanley Miller in the laboratory of Harold Urey. Miller subjected gaseous mixtures such as those presumed to exist on the prebiotic Earth, including NH_3, CH_4, H_2O, and H_2, to electrical sparks produced across a pair of electrodes (to simulate lightning) for periods of a week or more, then analyzed the contents of the closed reaction vessel **(Fig. 1–34)**. The gas phase of the resulting mixture contained CO and CO_2 as well as the starting materials. The water phase contained a variety of organic compounds, including some amino acids, hydroxy acids, aldehydes, and hydrogen cyanide (HCN).



(a)



(b)

FIGURE 1–34 Abiotic production of biomolecules. (a) Spark-discharge apparatus of the type used by Miller and Urey in experiments demonstrating abiotic formation of organic compounds under primitive atmospheric conditions. After subjection of the gaseous contents of the system to electrical sparks, products were collected by condensation. Biomolecules such as amino acids were among the products. **(b)** Stanley L. Miller (1930–2007) using his spark-discharge apparatus.

This experiment established the possibility of abiotic production of biomolecules in relatively short times under relatively mild conditions. When Miller's carefully stored samples were rediscovered in 2010 and examined with much more sensitive and discriminating techniques (high-performance liquid chromatography and mass spectrometry), his original observations were confirmed and greatly broadened. Previously unpublished experiments by Miller that included H_2S in the gas mixture (mimicking the "smoking" volcanic plumes at the sea bottom; **Fig. 1–35**) showed the formation of 23 amino acids and 7 organosulfur compounds, as well as a large number of other simple compounds that might have served as building blocks in prebiotic evolution.

More-refined laboratory experiments have provided good evidence that many of the chemical components of living cells, including polypeptides and RNA-like molecules, can form under these conditions. Polymers of RNA can act as catalysts in biologically significant reactions (see Chapters 26 and 27), and RNA probably played a crucial role in prebiotic evolution, both as catalyst and as information repository.

RNA or Related Precursors May Have Been the First Genes and Catalysts

In modern organisms, nucleic acids encode the genetic information that specifies the structure of enzymes, and enzymes catalyze the replication and repair of nucleic acids. The mutual dependence of these two classes of biomolecules brings up the perplexing question: which came first, DNA or protein?

The answer may be that they appeared about the same time, and RNA preceded them both. The discovery that RNA molecules can act as catalysts in their own formation suggests that RNA or a similar molecule may have been the first gene *and* the first catalyst. According to this scenario **(Fig. 1–36)**, one of the earliest stages of biological evolution was the chance formation of an RNA molecule that could catalyze the formation of other RNA molecules of the same sequence—a self-replicating, self-perpetuating RNA. The concentration of a self-replicating RNA molecule would increase exponentially, as one molecule formed several, several formed many, and so on. The fidelity of self-replication was presumably less than perfect, so the process would generate variants of the RNA, some of which might be even better able to self-replicate. In the competition for nucleotides, the most efficient of the self-replicating sequences would win, and less efficient replicators would fade from the population.

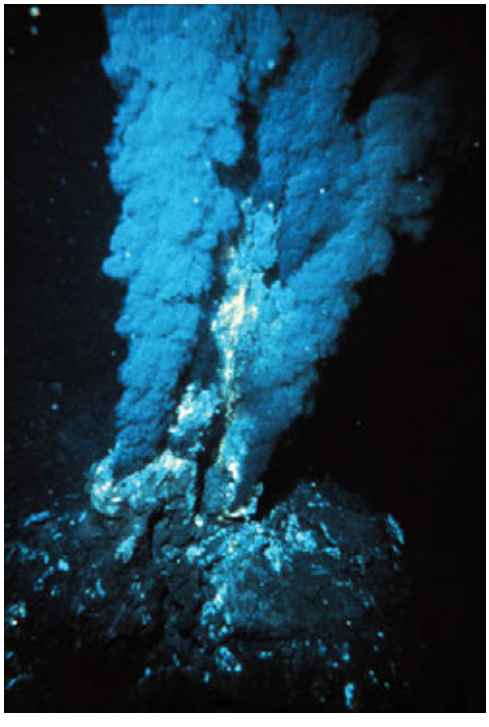


FIGURE 1–35 Black smokers. Hydrothermal vents in the sea floor emit superheated water rich in dissolved minerals. Black "smoke" is formed when the vented solution meets cold sea water and dissolved sulfides precipitate. Diverse life forms, including a variety of archaea and some remarkably complex multicellular animals, are found in the immediate vicinity of such vents, which may have been the sites of early biogenesis.

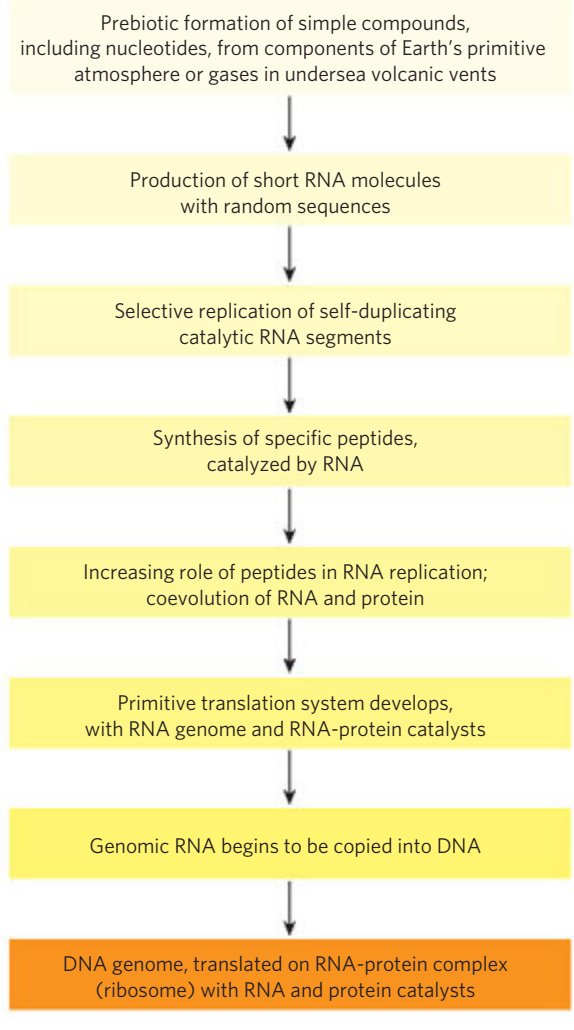


FIGURE 1–36 A possible "RNA world" scenario.

The division of function between DNA (genetic information storage) and protein (catalysis) was, according to the "RNA world" hypothesis, a later development. New variants of self-replicating RNA molecules developed, with the additional ability to catalyze the condensation of amino acids into peptides. Occasionally, the peptide(s) thus formed would reinforce the self-replicating ability of the RNA, and the pair—RNA molecule and helping peptide—could undergo further modifications in sequence, generating increasingly efficient self-replicating systems. The remarkable discovery that in the protein-synthesizing machinery of modern cells (ribosomes), RNA molecules, not proteins, catalyze the formation of peptide bonds is consistent with the RNA world hypothesis.

Some time after the evolution of this primitive protein-synthesizing system, there was a further development: DNA molecules with sequences complementary to the self-replicating RNA molecules took over the function of conserving the "genetic" information, and RNA molecules evolved to play roles in protein synthesis. (We explain in Chapter 8 why DNA is a more stable molecule than RNA and thus a better repository of inheritable information.) Proteins proved to be versatile catalysts and, over time, took over most of that function. Lipidlike compounds in the primordial mixture formed relatively impermeable layers around self-replicating collections of molecules. The concentration of proteins and nucleic acids within these lipid enclosures favored the molecular interactions required in self-replication.

The RNA world scenario is intellectually satisfying, but it leaves unanswered a vexing question: where did the nucleotides needed to make the initial RNA molecules come from? An alternative to this RNA world scenario supposes that simple metabolic pathways evolved first, perhaps at the hot vents in the ocean floor. A set of linked chemical reactions there might have produced precursors, including nucleotides, before the advent of lipid membranes or RNA. Without more experimental evidence, neither of these hypotheses can be disproved.

Biological Evolution Began More Than Three and a Half Billion Years Ago

Earth was formed about 4.6 billion years ago, and the first evidence of life dates to more than 3.5 billion years ago. In 1996, scientists working in Greenland found chemical evidence of life ("fossil molecules") from as far back as 3.85 billion years ago, forms of carbon embedded in rock that seem to have a distinctly biological origin. Somewhere on Earth during its first billion years the first simple organism arose, capable of replicating its own structure from a template (RNA?) that was the first genetic material. Because the terrestrial atmosphere at the dawn of life was nearly devoid of oxygen, and because there were few microorganisms to scavenge organic compounds formed by natural processes, these compounds were relatively stable. Given this stability and eons of time, the improbable became inevitable: lipid vesicles containing organic compounds and self-replicating RNA gave rise to the first cells (protocells), and those protocells with the greatest capacity for self-replication became more numerous. The process of biological evolution had begun.

The First Cell Probably Used Inorganic Fuels

The earliest cells arose in a reducing atmosphere (there was no oxygen) and probably obtained energy from inorganic fuels, such as ferrous sulfide and ferrous carbonate, both abundant on the early Earth. For example, the reaction

$$FeS + H_2S \longrightarrow FeS_2 + H_2$$

yields enough energy to drive the synthesis of ATP or similar compounds. The organic compounds these early cells required may have arisen by the nonbiological actions of lightning or of heat from volcanoes or thermal vents in the sea on components of the early atmosphere: CO, CO_2, N_2, NH_3, CH_4, and such. An alternative source of organic compounds has been proposed: extraterrestrial space. In 2006, the Stardust space mission brought back tiny particles of dust from the tail of a comet; the dust contained a variety of organic compounds, including the simple amino acid glycine.

Early unicellular organisms gradually acquired the ability to derive energy from compounds in their environment and to use that energy to synthesize more of their own precursor molecules, thereby becoming less dependent on outside sources. A very significant evolutionary event was the development of pigments capable of capturing the energy of light from the sun, which could be used to reduce, or "fix," CO_2 to form more complex, organic compounds. The original electron donor for these **photosynthetic** processes was probably H_2S, yielding elemental sulfur or sulfate (SO_4^{2-}) as the byproduct; later cells developed the enzymatic capacity to use H_2O as the electron donor in photosynthetic reactions, producing O_2 as waste. Cyanobacteria are the modern descendants of these early photosynthetic oxygen-producers.

Because the atmosphere of Earth in the earliest stages of biological evolution was nearly devoid of oxygen, the earliest cells were anaerobic. Under these conditions, chemotrophs could oxidize organic compounds to CO_2 by passing electrons not to O_2 but to acceptors such as SO_4^{2-}, in this case yielding H_2S as the product. With the rise of O_2-producing photosynthetic bacteria, the atmosphere became progressively richer in oxygen—a powerful oxidant and deadly poison to anaerobes. Responding to the evolutionary pressure of what Lynn Margulis and Dorion Sagan called the "oxygen holocaust," some lineages of microorganisms gave rise to aerobes that obtained energy by passing electrons from fuel molecules to oxygen.

Because the transfer of electrons from organic molecules to O_2 releases a great deal of energy, aerobic organisms had an energetic advantage over their anaerobic counterparts when both competed in an environment containing oxygen. This advantage translated into the predominance of aerobic organisms in O_2-rich environments.

Modern bacteria and archaea inhabit almost every ecological niche in the biosphere, and there are organisms capable of using virtually every type of organic compound as a source of carbon and energy. Photosynthetic microbes in both fresh and marine waters trap solar energy and use it to generate carbohydrates and all other cell constituents, which are in turn used as food by other forms of life. The process of evolution continues—and, in rapidly reproducing bacterial cells, on a time scale that allows us to witness it in the laboratory. One interesting line of research into evolutionary mechanisms aims at producing a "synthetic" cell in the laboratory (one in which the experimenter has provided every component from known, purified components). The first step in this direction involves determining the minimum number of genes necessary for life by examining the genomes of the simplest bacteria. The smallest known genome for a free-living bacterium is that of *Mycobacterium genitalium*, which comprises 580,000 base pairs encoding 483 genes. In 2010, scientists at the Craig Venter Institute succeeded in synthesizing the full chromosome of the mycobacterium in vitro, then incorporating that synthetic chromosome into a living bacterial cell of another species, which thereby acquired the properties of *Mycobacterium genitalium*. This technology opens the way to producing a synthetic cell, with the bare minimum of genes essential to life. With such a cell, one could hope to study in the laboratory the evolutionary processes by which protocells gradually diversified and became more complex.

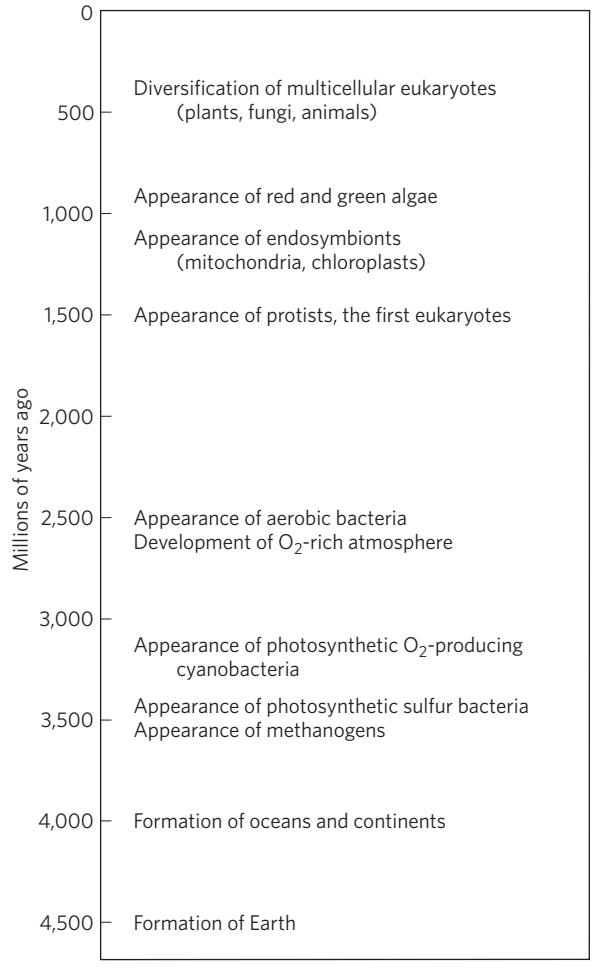


FIGURE 1–37 **Landmarks in the evolution of life on Earth.**

Eukaryotic Cells Evolved from Simpler Precursors in Several Stages

Starting about 1.5 billion years ago, the fossil record begins to show evidence of larger and more complex organisms, probably the earliest eukaryotic cells **(Fig. 1–37)**. Details of the evolutionary path from non-nucleated to nucleated cells cannot be deduced from the fossil record alone, but morphological and biochemical comparisons of modern organisms have suggested a sequence of events consistent with the fossil evidence.

Three major changes must have occurred. First, as cells acquired more DNA, the mechanisms required to fold it compactly into discrete complexes with specific proteins and to divide it equally between daughter cells at cell division became more elaborate. Specialized proteins were required to stabilize folded DNA and to pull the resulting DNA-protein complexes (*chromosomes*) apart during cell division. Second, as cells became larger, a system of intracellular membranes developed, including a double membrane surrounding the DNA. This membrane segregated the nuclear process of RNA synthesis on a DNA template from the cytoplasmic process of protein synthesis on ribosomes. Finally, according to a now widely accepted hypothesis advanced (initially, to much resistance) by Lynn Margulis, early eukaryotic cells, which were incapable of photosynthesis or aerobic metabolism, enveloped aerobic bacteria or photosynthetic bacteria to form **endosymbiotic** associations that eventually became permanent **(Fig. 1–38)**. Some aerobic bacteria evolved into the mitochondria of modern eukaryotes, and some photosynthetic cyanobacteria became the plastids, such as the chloroplasts of green algae, the likely ancestors of modern plant cells.

Lynn Margulis, 1938–2011

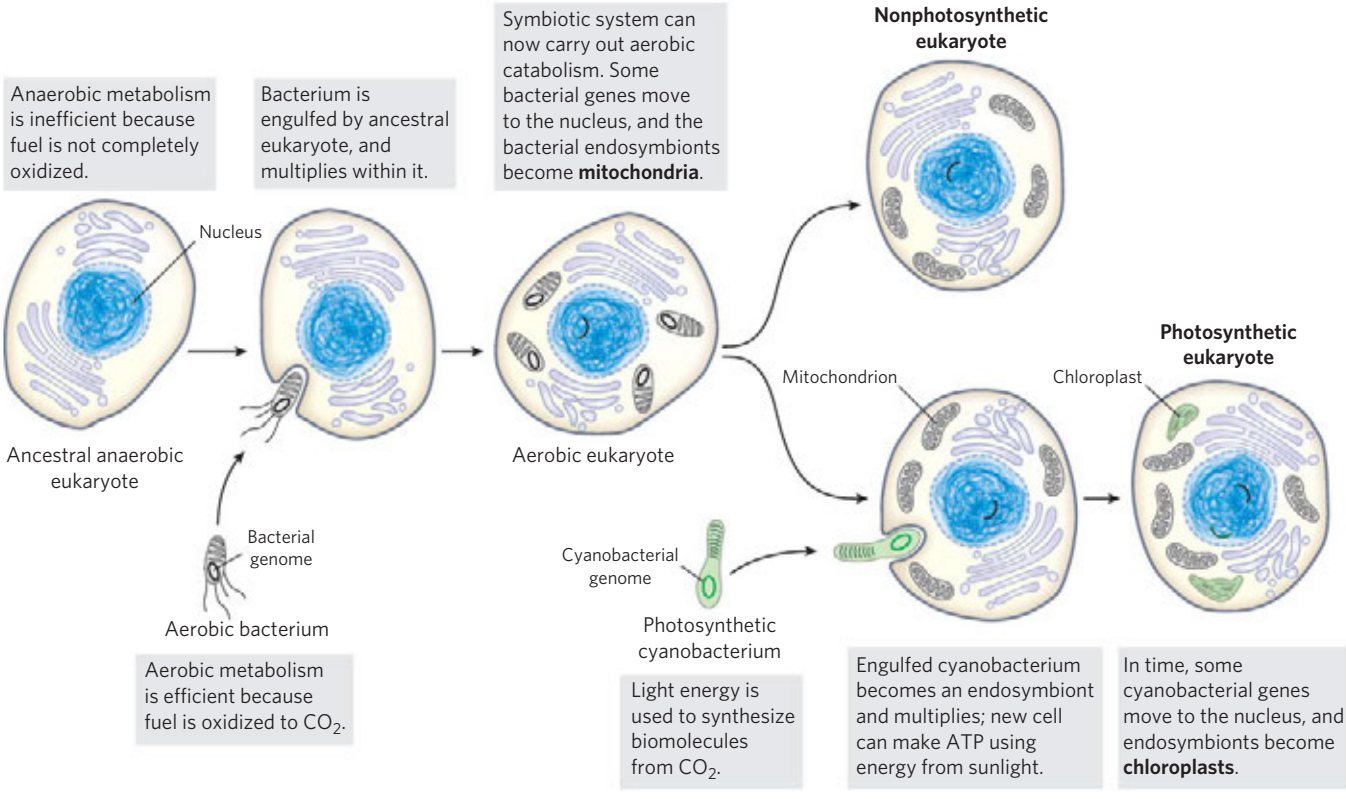


FIGURE 1–38 Evolution of eukaryotes through endosymbiosis. The earliest eukaryote, an anaerobe, acquired endosymbiotic purple bacteria, which carried with them their capacity for aerobic catabolism and became, over time, mitochondria. When photosynthetic cyanobacteria subsequently became endosymbionts of some aerobic eukaryotes, these cells became the photosynthetic precursors of modern green algae and plants.

At some later stage of evolution, unicellular organisms found it advantageous to cluster together, thereby acquiring greater motility, efficiency, or reproductive success than their free-living single-celled competitors. Further evolution of such clustered organisms led to permanent associations among individual cells and eventually to specialization within the colony—to cellular differentiation.

The advantages of cellular specialization led to the evolution of increasingly complex and highly differentiated organisms, in which some cells carried out the sensory functions, others the digestive, photosynthetic, or reproductive functions, and so forth. Many modern multicellular organisms contain hundreds of different cell types, each specialized for a function that supports the entire organism. Fundamental mechanisms that evolved early have been further refined and embellished through evolution. The same basic structures and mechanisms that underlie the beating motion of cilia in *Paramecium* and of flagella in *Chlamydomonas* are employed by the highly differentiated vertebrate sperm cell, for example.

Molecular Anatomy Reveals Evolutionary Relationships

Biochemists now have an enormously rich, ever increasing treasury of information on the molecular anatomy of cells that they can use to analyze evolutionary relationships and refine evolutionary theory. The sequence of the **genome**, the complete genetic endowment of an organism, has been determined for hundreds of bacteria and more than 40 archaea and for growing numbers of eukaryotic microorganisms, including *Saccharomyces cerevisiae* and *Plasmodium* species; plants, including *Arabidopsis thaliana* and rice; and multicellular animals, including *Caenorhabditis elegans* (a roundworm), *Drosophila melanogaster* (the fruit fly), mouse, rat, dog, chimpanzee, and *Homo sapiens* (Table 1–2). With such sequences in hand, detailed and quantitative comparisons among species can provide deep insight into the evolutionary process. Thus far, the molecular phylogeny derived from gene sequences is consistent with, but in many cases more precise than, the classical phylogeny based on macroscopic structures. Although organisms have continuously diverged at the level of gross anatomy, at the molecular level the basic unity of life is readily apparent; molecular structures and mechanisms are remarkably similar from the simplest to the most complex organisms. These similarities are most easily seen at the level of sequences, either the DNA sequences that encode proteins or the protein sequences themselves.

When two genes share readily detectable sequence similarities (nucleotide sequence in DNA or amino acid

TABLE 1–2 A Few of the Many Organisms Whose Genomes Have Been Completely Sequenced

Organism	Genome size (nucleotide pairs)	Biological interest
Mycoplasma genitalium	5.8×10^5	Smallest true organism
Helicobacter pylori	1.6×10^6	Causes gastric ulcers
Methanocaldococcus jannaschii	1.7×10^6	Archaeon; grows at 85 °C
Haemophilus influenzae	1.9×10^6	Causes bacterial influenza
Synechocystis sp.	3.9×10^6	Cyanobacterium
Bacillus subtilis	4.2×10^6	Common soil bacterium
Escherichia coli	4.6×10^6	Some strains are human pathogens
Saccharomyces cerevisiae	1.2×10^7	Unicellular eukaryote
Caenorhabditis elegans	1.0×10^8	Multicellular roundworm
Arabidopsis thaliana	1.2×10^8	Model plant
Drosophila melanogaster	1.8×10^8	Laboratory fly ("fruit fly")
Mus musculus	2.7×10^9	Laboratory mouse
Homo sapiens	3.0×10^9	Human

Source: www.ncbi.nlm.nih.gov/genome.

sequence in the proteins they encode), their sequences are said to be homologous and the proteins they encode are **homologs**. If two homologous genes occur in the *same* species, they are said to be paralogous and their protein products are **paralogs**. Paralogous genes are presumed to have been derived by gene duplication followed by gradual changes in the sequences of both copies. Typically, paralogous proteins are similar not only in sequence but also in three-dimensional structure, although they commonly have acquired different functions during their evolution.

Two homologous genes (or proteins) found in *different* species are said to be orthologous, and their protein products are **orthologs**. Orthologs are commonly found to have the same function in both organisms, and when a newly sequenced gene in one species is found to be strongly orthologous with a gene in another, this gene is presumed to encode a protein with the same function in both species. By this means, the function of gene products (proteins or RNA molecules) can be deduced from the genomic sequence, without any biochemical characterization of the molecules themselves. An **annotated genome** includes, in addition to the DNA sequence itself, a description of the likely function of each gene product, deduced from comparisons with other genomic sequences and established protein functions. Sometimes, by identifying the pathways (sets of enzymes) encoded in a genome, we can deduce from the genomic sequence alone the organism's metabolic capabilities.

The sequence differences between homologous genes may be taken as a rough measure of the degree to which the two species have diverged during evolution—of how long ago their common evolutionary precursor gave rise to two lines with different evolutionary fates. The larger the number of sequence differences, the earlier the divergence in evolutionary history. One can construct a phylogeny (family tree) in which the evolutionary distance between any two species is represented by their proximity on the tree (Fig. 1–4 is an example).

In the course of evolution, new structures, processes, or regulatory mechanisms are acquired, reflections of the changing genomes of the evolving organisms. The genome of a simple eukaryote such as yeast should have genes related to formation of the nuclear membrane, genes not present in bacteria or archaea. The genome of an insect should contain genes that encode proteins involved in specifying insects' characteristic segmented body plan, genes not present in yeast. The genomes of all vertebrate animals should share genes that specify the development of a spinal column, and those of mammals should have unique genes necessary for the development of the placenta, a characteristic of mammals—and so on. Comparisons of the whole genomes of species in each phylum are leading to the identification of genes critical to fundamental evolutionary changes in body plan and development.

Functional Genomics Shows the Allocations of Genes to Specific Cellular Processes

When the sequence of a genome is fully determined and each gene is assigned a function, molecular geneticists can group genes according to the processes (DNA

synthesis, protein synthesis, generation of ATP, and so forth) in which they function and thus find what fraction of the genome is allocated to each of a cell's activities. The largest category of genes in *E. coli*, *A. thaliana*, and *H. sapiens* consists of genes of (as yet) unknown function, which make up more than 40% of the genes in each species. The transporters that move ions and small molecules across plasma membranes take up a significant proportion of the genes in all three species, more in the bacterium and plant than in the mammal (10% of the ~4,400 genes of *E. coli*, ~8% of the ~32,000 genes of *A. thaliana*, and ~4% of the ~25,000 genes of *H. sapiens*). Genes that encode the proteins and RNA required for protein synthesis make up 3% to 4% of the *E. coli* genome, but in the more complex cells of *A. thaliana*, more genes are needed for targeting proteins to their final location in the cell than are needed to synthesize those proteins (about 6% and 2% of the genome, respectively). In general, the more complex the organism, the greater the proportion of its genome that encodes genes involved in the *regulation* of cellular processes and the smaller the proportion dedicated to the basic processes themselves, such as ATP generation and protein synthesis.

Genomic Comparisons Have Increasing Importance in Human Biology and Medicine

The genomes of chimpanzees and humans are 99.9% identical, yet the differences between the two species are vast. The relatively few differences in genetic endowment must explain the possession of language by humans, the extraordinary athleticism of chimpanzees, and myriad other differences. Genomic comparison is allowing researchers to identify candidate genes linked to divergences in the developmental programs of humans and the other primates and to the emergence of complex functions such as language. The picture will become clearer only as more primate genomes become available for comparison with the human genome.

Similarly, the differences in genetic endowment among humans are vanishingly small compared with the differences between humans and chimpanzees, yet these differences account for the variety among us—including differences in health and in susceptibility to chronic diseases. We have much to learn about the variability in sequence among humans, and the availability of genomic information will almost certainly transform medical diagnosis and treatment. We may expect that for some genetic diseases, palliatives will be replaced by cures; and that for disease susceptibilities associated with particular genetic markers, forewarning and perhaps increased preventive measures will prevail. Today's "medical history" may be replaced by a "medical forecast." ■

SUMMARY 1.5 Evolutionary Foundations

- Occasional inheritable mutations yield organisms that are better suited for survival and reproduction in an ecological niche, and their progeny come to dominate the population in that niche. This process of mutation and selection is the basis for the Darwinian evolution that led from the first cell to all modern organisms. The large number of genes shared by all living organisms explains their fundamental similarities.
- Life originated about 3.5 billion years ago, most likely with the formation of a membrane-enclosed compartment containing a self-replicating RNA molecule. The components for the first cell may have been produced near thermal vents at the bottom of the sea or by the action of lightning and high temperature on simple atmospheric molecules such as CO_2 and NH_3.
- The catalytic and genetic roles played by the early RNA genome were, over time, taken over by proteins and DNA, respectively.
- Eukaryotic cells acquired the capacity for photosynthesis and oxidative phosphorylation from endosymbiotic bacteria. In multicellular organisms, differentiated cell types specialize in one or more of the functions essential to the organism's survival.
- Knowledge of the complete genomic nucleotide sequences of organisms from different branches of the phylogenetic tree provides insights into evolution and offers great opportunities in human medicine.

Key Terms

All terms are defined in the glossary.

metabolite 3
nucleus 3
genome 3
eukaryotes 3
bacteria 4
archaea 4
cytoskeleton 8
stereoisomers 16
configuration 16
chiral center 17
conformation 19
entropy, *S* 23
enthalpy, *H* 23
free-energy change, ΔG 23
endergonic reaction 23
exergonic reaction 23
equilibrium 25
standard free-energy change, $\Delta G°$ 26
activation energy, $\Delta G^{‡}$ 27
catabolism 28
anabolism 28
metabolism 28
systems biology 29
mutation 32

Further Reading

General

Fruton, J.S. (1999) *Proteins, Enzymes, Genes: The Interplay of Chemistry and Biochemistry*, Yale University Press, New Haven.

A distinguished historian of biochemistry traces the development of this science and discusses its impact on medicine, pharmacy, and agriculture.

Harold, F.M. (2001) *The Way of the Cell: Molecules, Organisms, and the Order of Life*, Oxford University Press, Oxford.

Judson, H.F. (1996) *The Eighth Day of Creation: The Makers of the Revolution in Biology*, expanded edn, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

A highly readable and authoritative account of the rise of biochemistry and molecular biology in the twentieth century.

Kornberg, A. (1987) The two cultures: chemistry and biology. *Biochemistry* **26**, 6888–6891.

The importance of applying chemical tools to biological problems, described by an eminent practitioner.

Monod, J. (1971) *Chance and Necessity*, Alfred A. Knopf, Inc., New York. [Paperback edition, Vintage Books, 1972.] Originally published (1970) as *Le hasard et la nécessité*, Editions du Seuil, Paris.

An exploration of the philosophical implications of biological knowledge.

Morowitz, H.J. (2002) *The Emergence of Everything (How the World Became Complex)*, Oxford University Press, Oxford.

Short, beautifully written discussion of the emergence of complex organisms from simple beginnings.

Pace, N.R. (2001) The universal nature of biochemistry. *Proc. Natl. Acad. Sci. USA* **98**, 805–808.

A short discussion of the minimal definition of life, on Earth and elsewhere.

Cellular Foundations

Hardin, J., Bertoni, G.P., & Kleinsmith, L.J. (2011) *Becker's World of the Cell*, 8th edn, The Benjamin/Cummings Publishing Company, Redwood City, CA.

An excellent introductory textbook of cell biology.

Hosking, C.R. & Schwartz, J.L. (2009) The future's bright: imaging cell biology in the 21st century. *Trends Cell Biol.* **19**, 553–554.

Short editorial introduction to this whole issue of *Trends in Cell Biology*, about new methods for imaging cells.

Lodish, H., Berk, A., Kaiser, C.A., Krieger, M., Scott, M.R., Bretscher, A., Ploegh, H., & Amon, A. (2012) *Molecular Cell Biology*, 7th edn, W. H. Freeman and Company, New York.

A superb text, useful for this and later chapters.

Sadava, D., Hillis, D.M., Heller, H.C., & Berenbaum, M. (2010) *Life: The Science of Biology*, 9th edn, W. H. Freeman and Company, New York.

Wilson, C., Venditti, R., Rega, L.R., Colanzi, A., D'Angelo, G., & DeMatteis, M.A. (2011) The Golgi apparatus: an organelle with multiple complex functions. *Biochem. J.* **433**, 1–9.

Excellent intermediate-level review of the roles of the Golgi complex.

Chemical Foundations

Barta, N.S. & Stille, J.R. (1994) Grasping the concepts of stereochemistry. *J. Chem. Educ.* **71**, 20–23.

A clear description of the RS system for naming stereoisomers, with practical suggestions for determining and remembering the "handedness" of isomers.

Thall, E. (1996) When drug molecules look in the mirror. *J. Chem. Educ.* **73**, 481–484.

Biological difference in (*R*) and (*S*) isomers of drugs.

Vollhardt, K.P.C. & Shore, N.E. (2011) *Organic Chemistry: Structure and Function*, 6th edn, W. H. Freeman and Company, New York.

Up-to-date discussions of stereochemistry, functional groups, reactivity, and the chemistry of the principal classes of biomolecules.

Physical Foundations

Atkins, P.W. & de Paula, J. (2012) *Physical Chemistry for the Life Sciences*, 2nd edn, W. H. Freeman and Company, New York.

Blum, H.F. (1968) *Time's Arrow and Evolution*, 3rd edn, Princeton University Press, Princeton.

An excellent discussion of the way the second law of thermodynamics has influenced biological evolution.

Genetic Foundations

Griffiths, A.J.F., Wessler, S.R., Lewinton, R.C. & Carroll, S. (2008) *An Introduction to Genetic Analysis*, 9th edn, W. H. Freeman and Company, New York.

Jacob, F. (1973) *The Logic of Life: A History of Heredity*, Pantheon Books, Inc., New York. Originally published (1970) as *La logique du vivant: une histoire de l'hérédité*, Editions Gallimard, Paris.

A fascinating historical and philosophical account of the route to our present molecular understanding of life.

Pierce, B. (2012) *Genetics: A Conceptual Approach*, 4th edn, W. H. Freeman and Company, New York.

Evolutionary Foundations

Adams, J.C. (2009) Molecular and cellular evolution: a celebration of the 200th anniversary of the birth of Charles Darwin. *Int. J. Biochem. Cell Biol.* **41**, 249.

An editorial introduction to 14 essays on various aspects of molecular and cellular evolution that make up this issue of the journal.

Brow, J.R. & Doolittle, W.F. (1997) Archaea and the prokaryote-to-eukaryote transition. *Microbiol. Mol. Biol. Rev.* **61**, 456–502.

A very thorough discussion of the arguments for placing the Archaea on the phylogenetic branch that led to multicellular organisms.

Budin, I. & Szostak, J.W. (2010) Expanding roles for diverse physical phenomena during the origin of life. *Annu. Rev. Biophys.* **39**, 45–63.

A thoughtful discussion of the physical factors that may have influenced the course of prebiotic evolution.

Carroll, S.B. (2006) *The Making of the Fittest: DNA and the Ultimate Forensic Record of Evolution*, W.W. Norton & Company, Inc., New York.

Cavicchioli, R. (2011) Archaea—timeline of the third domain. *Nat. Rev. Microbiol.* **9**, 51–61.

Intermediate-level description of the discovery and investigation of the Archaea.

de Duve, C. (1995) The beginnings of life on earth. *Am. Sci.* **83**, 428–437.

One scenario for the succession of chemical steps that led to the first living organism.

de Duve, C. (1996) The birth of complex cells. *Sci. Am.* **274** (April), 50–57.

Evolution of Catalytic Function. (1987) *Cold Spring Harb. Symp. Quant. Biol.* **52**.

A collection of almost 100 articles on all aspects of prebiotic and early biological evolution; excellent source on molecular evolution.

Gesteland, R.F., Atkins, J.F., & Cech, T.R. (eds). (2006) *The RNA World*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

A collection of stimulating reviews on a wide range of topics related to the RNA world scenario.

Knoll, A.H. (2003) *Life on a Young Planet: The First Three Billion Years of Evolution on Earth*, Princeton University Press, Princeton.

Discussion of the early stages of evolution, with emphasis on the geochemistry of the environment in which it occurred.

Lazcano, A. & Miller, S.L. (1996) The origin and early evolution of life: prebiotic chemistry, the pre-RNA world, and time. *Cell* **85**, 793–798.

Brief review of developments in studies of the origin of life: primitive atmospheres, submarine vents, autotrophic versus heterotrophic origin, the RNA and pre-RNA worlds, and the time required for life to arise.

Martin, W., Baross, J., Kelley, D., & Russell, M.J. (2008) Hydrothermal vents and the origin of life. *Nat. Rev. Microbiol.* **6**, 805–814.

Miller, S.L. (1987) Which organic compounds could have occurred on the prebiotic earth? *Cold Spring Harb. Symp. Quant. Biol.* **52**, 17–27.

Summary of laboratory experiments on chemical evolution, by the person who did the original Miller-Urey experiment.

Noireaux, V., Maeda, Y.T., & Libchaber, A. (2011) Development of an artificial cell, from self-organization to computation and self-reproduction. *Proc. Natl. Acad. Sci. USA* **108**, 3473–3480.

Interesting discussion of the steps that must be taken to produce a living cell from scratch.

Parker, E.T., Cleaves, H.J., Dworkin, J.P., Glavin, D.P., Callahan, M., Aubrey, A., Lazcano, A., & Bada, J.L. (2011) Primordial synthesis of amines and amino acids in a 1958 Miller H_2S-rich spark discharge experiment. *Proc. Natl. Acad. Sci. USA* **108**, 5526–5531.

A modern reexamination of samples generated in the 1958 experiments of Miller, using H_2S as one of the atmospheric gases.

Stiller, J.W. (2007) Plastid endosymbiosis, genome evolution and the origin of green plants. *Trends Plant Sci.* **12**, 391–396.

Woese, C.R. (2002) On the evolution of cells. *Proc. Natl. Acad. Sci. USA* **99**, 8742–8747.

Short, clear review.

Woese, C.R. (2004) A new biology for a new century. *Microbiol. Mol. Biol. Rev.* **68**, 173–186.

Development of current thinking about cellular evolution by one of the seminal thinkers in the field.

Woese, C.R., Kandler, O., & Wheelis, M.L. (1990) Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proc. Natl. Acad. Sci. USA* **87**, 4576–4579.

The arguments for dividing all living organisms into three domains.

Problems

Some problems related to the contents of the chapter follow. (In solving end-of-chapter problems, you may wish to refer to the tables on the inside of the back cover.) Each problem has a title for easy reference and discussion. For all numerical problems, keep in mind that answers should be expressed with the correct number of significant figures. Brief solutions are provided in Appendix B; expanded solutions are published in the *Absolute Ultimate Study Guide to Accompany Principles of Biochemistry.*

1. The Size of Cells and Their Components

(a) If you were to magnify a cell 10,000 fold (typical of the magnification achieved using an electron microscope), how big would it appear? Assume you are viewing a "typical" eukaryotic cell with a cellular diameter of 50 μm.

(b) If this cell were a muscle cell (myocyte), how many molecules of actin could it hold? Assume the cell is spherical and no other cellular components are present; actin molecules are spherical, with a diameter of 3.6 nm. (The volume of a sphere is $4/3\ \pi r^3$.)

(c) If this were a liver cell (hepatocyte) of the same dimensions, how many mitochondria could it hold? Assume the cell is spherical; no other cellular components are present; and the mitochondria are spherical, with a diameter of 1.5 μm.

(d) Glucose is the major energy-yielding nutrient for most cells. Assuming a cellular concentration of 1 mM (that is, 1 millimole/L), calculate how many molecules of glucose would be present in our hypothetical (and spherical) eukaryotic cell. (Avogadro's number, the number of molecules in 1 mol of a nonionized substance, is 6.02×10^{23}.)

(e) Hexokinase is an important enzyme in the metabolism of glucose. If the concentration of hexokinase in our eukaryotic cell is 20 μM, how many glucose molecules are present per hexokinase molecule?

2. Components of *E. coli* *E. coli* cells are rod-shaped, about 2 μm long and 0.8 μm in diameter. The volume of a cylinder is $\pi r^2 h$, where h is the height of the cylinder.

(a) If the average density of *E. coli* (mostly water) is 1.1×10^3 g/L, what is the mass of a single cell?

(b) *E. coli* has a protective cell envelope 10 nm thick. What percentage of the total volume of the bacterium does the cell envelope occupy?

(c) *E. coli* is capable of growing and multiplying rapidly because it contains some 15,000 spherical ribosomes (diameter 18 nm), which carry out protein synthesis. What percentage of the cell volume do the ribosomes occupy?

3. Genetic Information in *E. coli* DNA The genetic information contained in DNA consists of a linear sequence of coding units, known as codons. Each codon is a specific sequence of three deoxyribonucleotides (three deoxyribonucleotide pairs in double-stranded DNA), and each codon codes for a single amino acid unit in a protein. The molecular weight of an *E. coli* DNA molecule is about 3.1×10^9 g/mol. The average molecular weight of a nucleotide pair is 660 g/mol, and each nucleotide pair contributes 0.34 nm to the length of DNA.

(a) Calculate the length of an *E. coli* DNA molecule. Compare the length of the DNA molecule with the cell dimensions (see Problem 2). How does the DNA molecule fit into the cell?

(b) Assume that the average protein in *E. coli* consists of a chain of 400 amino acids. What is the maximum number of proteins that can be coded by an *E. coli* DNA molecule?

4. The High Rate of Bacterial Metabolism Bacterial cells have a much higher rate of metabolism than animal cells. Under ideal conditions some bacteria double in size and divide every 20 min, whereas most animal cells under rapid growth conditions require 24 hours. The high rate of bacterial metabolism requires a high ratio of surface area to cell volume.

(a) Why does surface-to-volume ratio affect the maximum rate of metabolism?

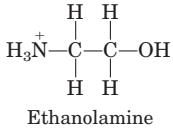
(b) Calculate the surface-to-volume ratio for the spherical bacterium *Neisseria gonorrhoeae* (diameter 0.5 μm), responsible for the disease gonorrhea. Compare it with the

surface-to-volume ratio for a globular amoeba, a large eukaryotic cell (diameter 150 μm). The surface area of a sphere is $4\pi r^2$.

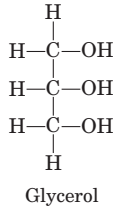
5. Fast Axonal Transport Neurons have long thin processes called axons, structures specialized for conducting signals throughout the organism's nervous system. Some axonal processes can be as long as 2 m—for example, the axons that originate in your spinal cord and terminate in the muscles of your toes. Small membrane-enclosed vesicles carrying materials essential to axonal function move along microtubules of the cytoskeleton, from the cell body to the tips of the axons. If the average velocity of a vesicle is 1 μm/s, how long does it take a vesicle to move from a cell body in the spinal cord to the axonal tip in the toes?

6. Is Synthetic Vitamin C as Good as the Natural Vitamin? A claim put forth by some purveyors of health foods is that vitamins obtained from natural sources are more healthful than those obtained by chemical synthesis. For example, pure L-ascorbic acid (vitamin C) extracted from rose hips is better than pure L-ascorbic acid manufactured in a chemical plant. Are the vitamins from the two sources different? Can the body distinguish a vitamin's source?

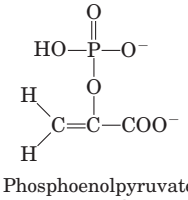
7. Identification of Functional Groups Figures 1–16 and 1–17 show some common functional groups of biomolecules. Because the properties and biological activities of biomolecules are largely determined by their functional groups, it is important to be able to identify them. In each of the compounds below, circle and identify by name each functional group.



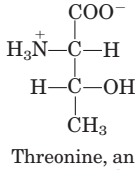
Ethanolamine
(a)



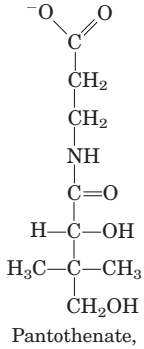
Glycerol
(b)



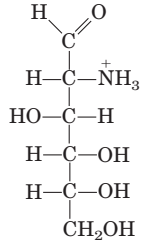
Phosphoenolpyruvate, an intermediate in glucose metabolism
(c)



Threonine, an amino acid
(d)

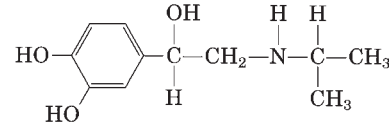


Pantothenate, a vitamin
(e)



D-Glucosamine
(f)

8. Drug Activity and Stereochemistry The quantitative differences in biological activity between the two enantiomers of a compound are sometimes quite large. For example, the D isomer of the drug isoproterenol, used to treat mild asthma, is 50 to 80 times more effective as a bronchodilator than the L isomer. Identify the chiral center in isoproterenol. Why do the two enantiomers have such radically different bioactivity?

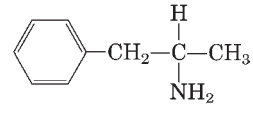


Isoproterenol

9. Separating Biomolecules In studying a particular biomolecule (a protein, nucleic acid, carbohydrate, or lipid) in the laboratory, the biochemist first needs to separate it from other biomolecules in the sample—that is, to *purify* it. Specific purification techniques are described later in the text. However, by looking at the monomeric subunits of a biomolecule, you should have some ideas about the characteristics of the molecule that would allow you to separate it from other molecules. For example, how would you separate (a) amino acids from fatty acids and (b) nucleotides from glucose?

10. Silicon-Based Life? Silicon is in the same group of the periodic table as carbon and, like carbon, can form up to four single bonds. Many science fiction stories have been based on the premise of silicon-based life. Is this realistic? What characteristics of silicon make it *less* well adapted than carbon as the central organizing element for life? To answer this question, consider what you have learned about carbon's bonding versatility, and refer to a beginning inorganic chemistry textbook for silicon's bonding properties.

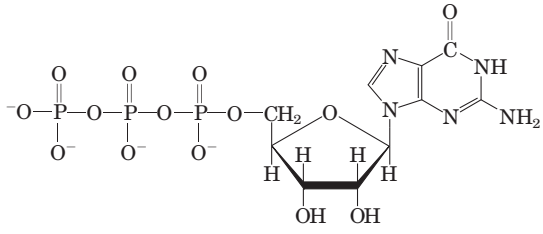
11. Drug Action and Shape of Molecules Some years ago two drug companies marketed a drug under the trade names Dexedrine and Benzedrine. The structure of the drug is shown below.



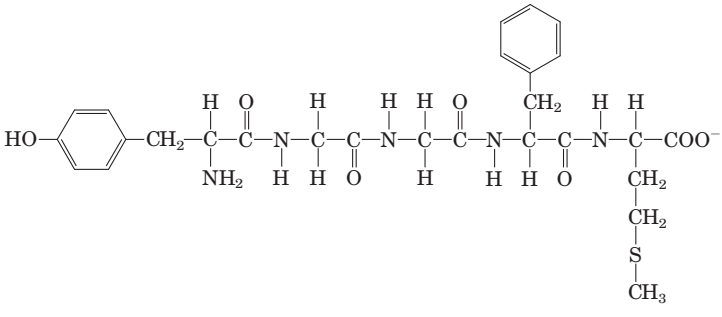
The physical properties (C, H, and N analysis, melting point, solubility, etc.) of Dexedrine and Benzedrine were identical. The recommended oral dosage of Dexedrine (which is still available) was 5 mg/day, but the recommended dosage of Benzedrine (no longer available) was twice that. Apparently it required considerably more Benzedrine than Dexedrine to yield the same physiological response. Explain this apparent contradiction.

12. Components of Complex Biomolecules Figure 1–10 shows the major components of complex biomolecules. For each of the three important biomolecules below (shown in their ionized forms at physiological pH), identify the constituents.

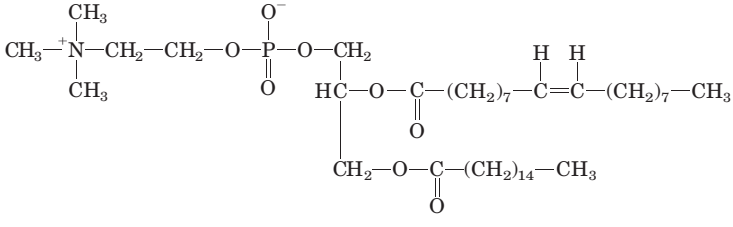
(a) Guanosine triphosphate (GTP), an energy-rich nucleotide that serves as a precursor to RNA:



(b) Methionine enkephalin, the brain's own opiate:



(c) Phosphatidylcholine, a component of many membranes:



13. Determination of the Structure of a Biomolecule An unknown substance, X, was isolated from rabbit muscle. Its structure was determined from the following observations and experiments. Qualitative analysis showed that X was composed entirely of C, H, and O. A weighed sample of X was completely oxidized, and the H_2O and CO_2 produced were measured; this quantitative analysis revealed that X contained 40.00% C, 6.71% H, and 53.29% O by weight. The molecular mass of X, determined by mass spectrometry, was 90.00 u (atomic mass units; see Box 1–1). Infrared spectroscopy showed that X contained one double bond. X dissolved readily in water to give an acidic solution; the solution demonstrated optical activity when tested in a polarimeter.

(a) Determine the empirical and molecular formula of X.

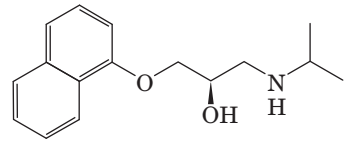
(b) Draw the possible structures of X that fit the molecular formula and contain one double bond. Consider *only* linear or branched structures and disregard cyclic structures. Note that oxygen makes very poor bonds to itself.

(c) What is the structural significance of the observed optical activity? Which structures in (b) are consistent with the observation?

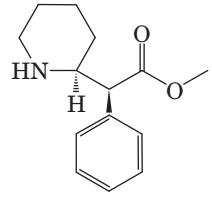
(d) What is the structural significance of the observation that a solution of X was acidic? Which structures in (b) are consistent with the observation?

(e) What is the structure of X? Is more than one structure consistent with all the data?

14. Naming Stereoisomers with One Chiral Carbon Using the RS System Propranolol is a chiral compound. (*R*)-Propranolol is used as a contraceptive; (*S*)-propranolol is used to treat hypertension. Identify the chiral carbon in the structure below. Is this the (*R*) or the (*S*) isomer? Draw the other isomer.



15. Naming Stereoisomers with Two Chiral Carbons Using the RS System The (*R,R*) isomer of methylphenidate (Ritalin) is used to treat attention deficit hyperactivity disorder (ADHD). The (*S,S*) isomer is an antidepressant. Identify the two chiral carbons in the structure below. Is this the (*R,R*) or the (*S,S*) isomer? Draw the other isomer.



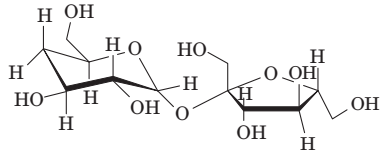
Data Analysis Problem

16. Interaction of Sweet-Tasting Molecules with Taste Receptors Many compounds taste sweet to humans. Sweet taste results when a molecule binds to the sweet receptor, one type of taste receptor, on the surface of certain tongue cells. The stronger the binding, the lower the concentration required to saturate the receptor and the sweeter a given concentration of that substance tastes. The standard free-energy change, $\Delta G°$, of the binding reaction between a sweet molecule and a sweet receptor can be measured in kilojoules or kilocalories per mole.

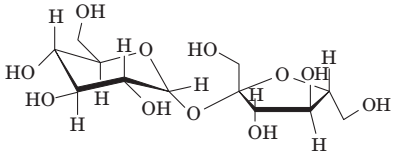
Sweet taste can be quantified in units of "molar relative sweetness" (MRS), a measure that compares the sweetness of a substance to the sweetness of sucrose. For example, saccharin has an MRS of 161; this means that saccharin is 161 times sweeter than sucrose. In practical terms, this is measured by asking human subjects to compare the sweetness of solutions containing different concentrations of each compound. Sucrose and saccharin taste equally sweet when sucrose is at a concentration 161 times higher than that of saccharin.

(a) What is the relationship between MRS and the $\Delta G°$ of the binding reaction? Specifically, would a more negative $\Delta G°$ correspond to a higher or lower MRS? Explain your reasoning.

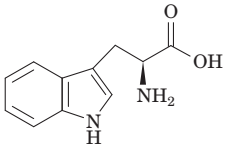
Shown on the next page are the structures of 10 compounds, all of which taste sweet to humans. The MRS and $\Delta G°$ for binding to the sweet receptor are given for each substance.



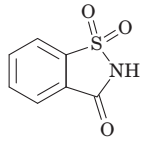
Deoxysucrose
MRS = 0.95
$\Delta G^\circ = -6.67$ kcal/mol



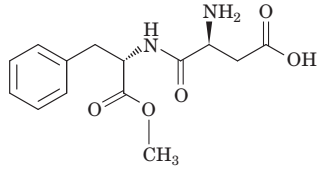
Sucrose
MRS = 1
$\Delta G^\circ = -6.71$ kcal/mol



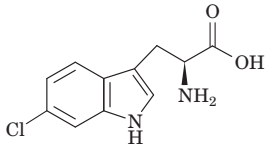
D-Tryptophan
MRS = 21
$\Delta G^\circ = -8.5$ kcal/mol



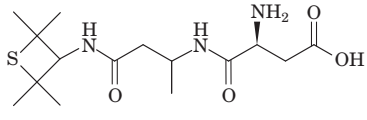
Saccharin
MRS = 161
$\Delta G^\circ = -9.7$ kcal/mol



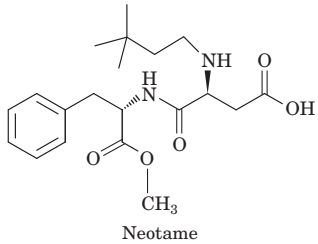
Aspartame
MRS = 172
$\Delta G^\circ = -9.7$ kcal/mol



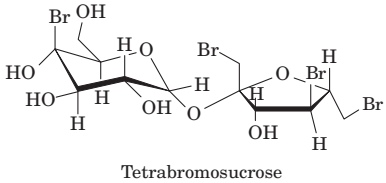
6-Chloro-D-tryptophan
MRS = 906
$\Delta G^\circ = -10.7$ kcal/mol



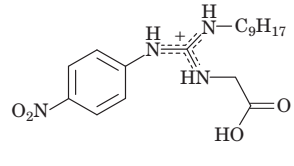
Alitame
MRS = 1,937
$\Delta G^\circ = -11.1$ kcal/mol



Neotame
MRS = 11,057
$\Delta G^\circ = -12.1$ kcal/mol



Tetrabromosucrose
MRS = 13,012
$\Delta G^\circ = -12.2$ kcal/mol



Sucronic acid
MRS = 200,000
$\Delta G^\circ = -13.8$ kcal/mol

Morini, Bassoli, and Temussi (2005) used computer-based methods (often referred to as "in silico" methods) to model the binding of sweet molecules to the sweet receptor.

(b) Why is it useful to have a computer model to predict the sweetness of molecules, instead of a human- or animal-based taste assay?

In earlier work, Schallenberger and Acree (1967) had suggested that all sweet molecules include an "AH-B" structural group, in which "A and B are electronegative atoms separated by a distance of greater than 2.5 Å [0.25 nm] but less than 4 Å [0.4 nm]. H is a hydrogen atom attached to one of the electronegative atoms by a covalent bond."

(c) Given that the length of a "typical" single bond is about 0.15 nm, identify the AH-B group(s) in each of the molecules shown on the left.

(d) Based on your findings from (c), give two objections to the statement that "molecules containing an AH-B structure will taste sweet."

(e) For two of the molecules shown here, the AH-B model *can* be used to explain the difference in MRS and ΔG°. Which two molecules are these, and how would you use them to support the AH-B model?

(f) Several of the molecules have closely related structures but very different MRS and ΔG° values. Give two such examples, and use these to argue that the AH-B model is unable to explain the observed differences in sweetness.

In their computer-modeling study, Morini and coauthors used the three-dimensional structure of the sweet receptor and a molecular dynamics modeling program called GRAMM to predict the ΔG° of binding of sweet molecules to the sweet receptor. First, they "trained" their model—that is, they refined the parameters so that the ΔG° values predicted by the model matched the known ΔG° values for one set of sweet molecules (the "training set"). They then "tested" the model by asking it to predict the ΔG° values for a new set of molecules (the "test set").

(g) Why did Morini and colleagues need to test their model against a different set of molecules from the set it was trained on?

(h) The researchers found that the predicted ΔG° values for the test set differed from the actual values by, on average, 1.3 kcal/mol. Using the values given with the molecular structures, estimate the resulting error in MRS values.

References

Morini, G., Bassoli, A., & Temussi, P.A. (2005) From small sweeteners to sweet proteins: anatomy of the binding sites of the human T1R2_T1R3 receptor. *J. Med. Chem.* **48**, 5520–5529.

Schallenberger, R.S. & Acree, T.E. (1967) Molecular theory of sweet taste. *Nature* **216**, 480–482.