32nd Edition

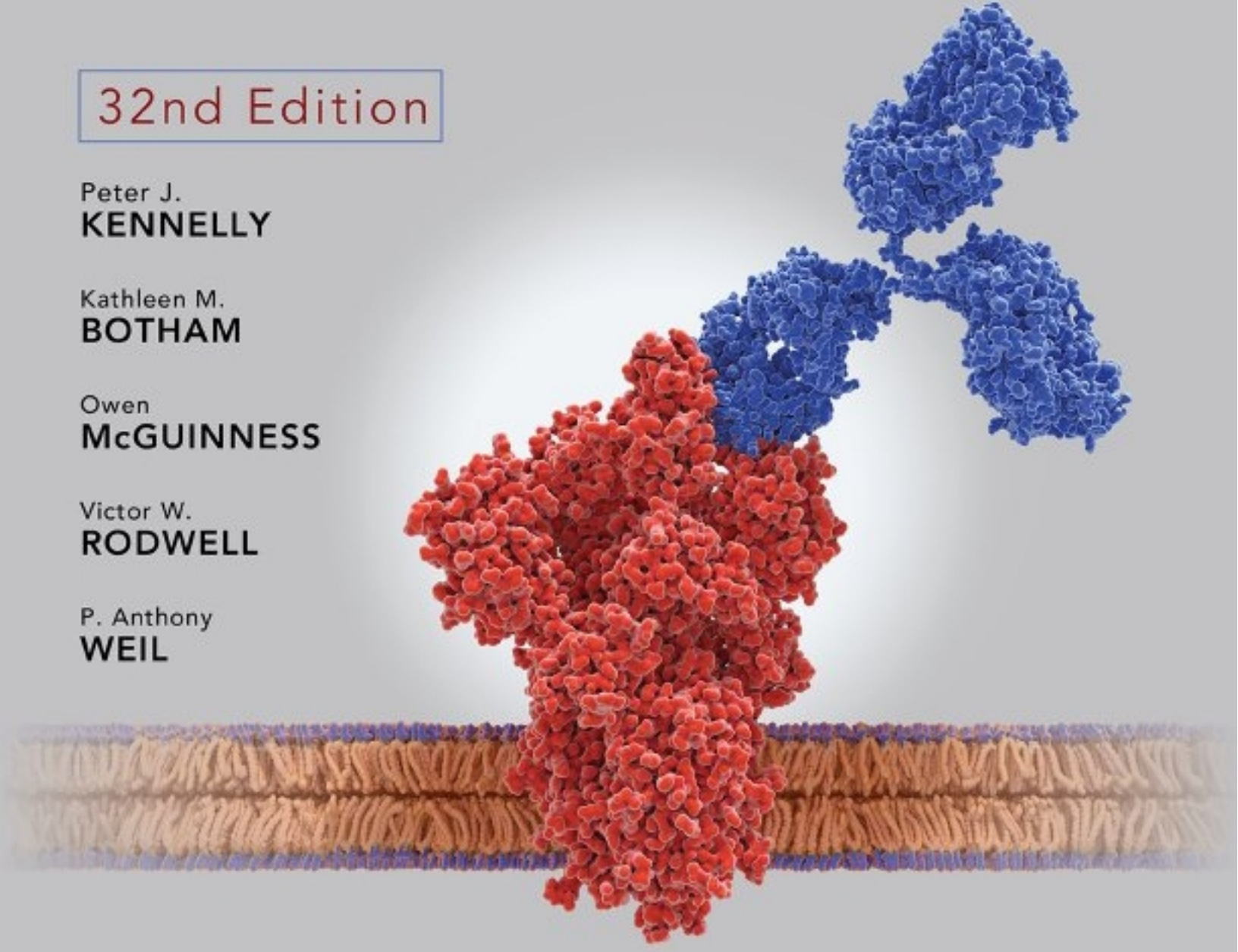
Peter J.
KENNELLY

Kathleen M.
BOTHAM

Owen
McGUINNESS

Victor W.
RODWELL

P. Anthony
WEIL



HARPER'S ILLUSTRATED BIOCHEMISTRY

Mc Graw Hill LANGE®

a LANGE medical book

Harper's Illustrated Biochemistry

THIRTY-SECOND EDITION

Peter J. Kennelly, PhD
Professor
Department of Biochemistry
Virginia Tech
Blacksburg, Virginia

Kathleen M. Botham, PhD, DSc
Emeritus Professor of Biochemistry
Department of Comparative Biomedical Sciences
Royal Veterinary College
University of London
London, United Kingdom

Owen P. McGuinness, PhD
Professor
Department of Molecular Physiology & Biophysics
Vanderbilt University
School of Medicine
Nashville, Tennessee

Victor W. Rodwell, PhD
Professor (Emeritus) of Biochemistry
Purdue University
West Lafayette, Indiana

P. Anthony Weil, PhD
Professor Emeritus of Molecular Physiology & Biophysics
Vanderbilt University
Nashville, Tennessee

Mc Graw Hill

New York Chicago San Francisco Athens London Madrid Mexico City
Milan New Delhi Singapore Sydney Toronto

Copyright © 2023 by McGraw Hill, LLC. All rights reserved. Except as permitted under the United States Copyright Act of 1976, no part of this publication may be reproduced or distributed in any form or by any means, or stored in a database or retrieval system, without the prior written permission of the publisher.

ISBN: 978-1-26-046995-0
MHID: 1-26-046995-6

The material in this eBook also appears in the print version of this title: ISBN: 978-1-26-046994-3,
MHID: 1-26-046994-8.

eBook conversion by codeMantra
Version 1.0

All trademarks are trademarks of their respective owners. Rather than put a trademark symbol after every occurrence of a trademarked name, we use names in an editorial fashion only, and to the benefit of the trademark owner, with no intention of infringement of the trademark. Where such designations appear in this book, they have been printed with initial caps.

McGraw-Hill Education eBooks are available at special quantity discounts to use as premiums and sales promotions or for use in corporate training programs. To contact a representative, please visit the Contact Us page at www.mhprofessional.com.

Notice

Medicine is an ever-changing science. As new research and clinical experience broaden our knowledge, changes in treatment and drug therapy are required. The authors and the publisher of this work have checked with sources believed to be reliable in their efforts to provide information that is complete and generally in accord with the standards accepted at the time of publication. However, in view of the possibility of human error or changes in medical sciences, neither the authors nor the publisher nor any other party who has been involved in the preparation or publication of this work warrants that the information contained herein is in every respect accurate or complete, and they disclaim all responsibility for any errors or omissions or for the results obtained from use of the information contained in this work. Readers are encouraged to confirm the information contained herein with other sources. For example and in particular, readers are advised to check the product information sheet included in the package of each drug they plan to administer to be certain that the information contained in this work is accurate and that changes have not been made in the recommended dose or in the contraindications for administration. This recommendation is of particular importance in connection with new or infrequently used drugs.

TERMS OF USE

This is a copyrighted work and McGraw-Hill Education and its licensors reserve all rights in and to the work. Use of this work is subject to these terms. Except as permitted under the Copyright Act of 1976 and the right to store and retrieve one copy of the work, you may not decompile, disassemble, reverse engineer, reproduce, modify, create derivative works based upon, transmit, distribute, disseminate, sell, publish or sublicense the work or any part of it without McGraw-Hill Education's prior consent. You may use the work for your own noncommercial and personal use; any other use of the work is strictly prohibited. Your right to use the work may be terminated if you fail to comply with these terms.

THE WORK IS PROVIDED "AS IS." McGRAW-HILL EDUCATION AND ITS LICENSORS MAKE NO GUARANTEES OR WARRANTIES AS TO THE ACCURACY, ADEQUACY OR COMPLETENESS OF OR RESULTS TO BE OBTAINED FROM USING THE WORK, INCLUDING ANY INFORMATION THAT CAN BE ACCESSED THROUGH THE WORK VIA HYPERLINK OR OTHERWISE, AND EXPRESSLY DISCLAIM ANY WARRANTY, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. McGraw-Hill Education and its licensors do not warrant or guarantee that the functions contained in the work will meet your requirements or that its operation will be uninterrupted or error free. Neither McGraw-Hill Education nor its licensors shall be liable to you or anyone else for any inaccuracy, error or omission, regardless of cause, in the work or for any damages resulting therefrom. McGraw-Hill Education has no responsibility for the content of any information accessed through the work. Under no circumstances shall McGraw-Hill Education and/or its licensors be liable for any indirect, incidental, special, punitive, consequential or similar damages that result from the use of or inability to use the work, even if any of them has been advised of the possibility of such damages. This limitation of liability shall apply to any claim or cause whatsoever whether such claim or cause arises in contract, tort or otherwise.

Co-Authors

David A. Bender, PhD
Professor (Emeritus) of Nutritional Biochemistry
University College London
London, United Kingdom

Peter L. Gross, MD, MSc, FRCP(C)
Associate Professor
Department of Medicine
McMaster University
Hamilton, Ontario, Canada

January D. Haile, PhD
Associate Professor of Biochemistry and Molecular Biology
Centre College
Danville, Kentucky

Molly Jacob, MD, PhD, MNASc, FRCPath
Professor of Biochemistry
Christian Medical College
Vellore, Tamil Nadu, India

Peter A. Mayes, PhD, DSc
Emeritus Professor of Veterinary Biochemistry
Royal Veterinary College
University of London
London, United Kingdom

Robert K. Murray, MD, PhD
Emeritus Professor of Biochemistry
University of Toronto
Toronto, Ontario, Canada

Margaret L. Rand, PhD
Senior Associate Scientist
Division of Hematology/Oncology
Hospital for Sick Children, Toronto
Professor, Department of Biochemistry
University of Toronto, Toronto, Canada

Joe Varghese, MD, PhD
Professor and Head of Biochemistry
Christian Medical College
Vellore, Tamil Nadu, India

This page intentionally left blank

Contents

This page intentionally left blank

Preface

The authors and publishers are pleased to present the thirty-second edition of *Harper's Illustrated Biochemistry*. The first edition, entitled *Harper's Biochemistry*, was published in 1939 under the sole authorship of Dr Harold Harper at the University of California School of Medicine, San Francisco, California. Presently entitled *Harper's Illustrated Biochemistry*, the book continues, as originally intended, to provide a concise survey of aspects of biochemistry most relevant to the study of medicine. Various authors have contributed to subsequent editions of this medically oriented biochemistry text, which is now observing its 83rd year.

Cover Illustration for the Thirty-Second Edition

The global COVID-19 pandemic has provided a dramatic, face-to-face demonstration of both the power and limitations of molecular medicine and epidemiology. The rapid development of highly effective vaccines was made possible by the adaptation of novel RNA-based approaches in which the patient's immune response is activated via the endogenous expression of genetically-encoded antigens, rather than the physical injection of a non-infectious antigen. Utilizing the patient's own cells as the bioreactor for generating antigens, rather than some animal or culture, enabled scientists to use the self-amplifying capacity of polynucleotides to accelerate both the speed of vaccine development and subsequent large-scale manufacture. The illustration on the cover of the thirty-second edition depicts a neutralizing antibody, in blue, bound to the spike protein on the surface of the SARS-CoV-2 coronavirus, better known as COVID-19, which is shown in red. The epitope to which the antibody binds overlaps that at which the virus binds to the ACE-2 receptor, the membrane protein by which the pathogen recognizes, binds to, and subsequently invades human cells. Therapeutic antibodies thus protect by physically blocking association of the Spike protein with the ACE-2 receptor.

Changes in the Thirty-Second Edition

As always, *Harper's Illustrated Biochemistry* continues to emphasize the close relationship of biochemistry to the understanding of diseases, their pathology, and the practice of medicine. With the retirement of long-time contributor David A. Bender, Prof. Owen P. McGuinness of Vanderbilt University has joined as a new coauthor. In addition to the fresh perspectives and novel insights provided by Prof. McGuinness, the contents of most chapters have been updated and provide the reader with the most current and pertinent information. For example, in Chapter 6 the description of the Bohr effect's contributions to CO_2 transport and release from the lungs has been reorganized and expanded, while Chapter 9 has been updated and reorganized to include expanded coverage of zymogen activation in enzyme regulation.

Organization of the Book

All 58 chapters of the thirty-second edition place major emphasis on the medical relevance of biochemistry. Topics are organized under 11 major headings. In order to assist study and to facilitate retention of the contained information, Questions follow each Section. An Answer Bank follows Chapter 58.

Section I includes a brief history of biochemistry and emphasizes the interrelationships between biochemistry and medicine. Water and the importance of homeostasis of intracellular pH are reviewed, and the various orders of proteins structure are addressed.

Section II begins with a chapter on hemoglobin. The next four chapters address the mechanism of action, kinetics, metabolic regulation of enzymes, and the role of metal ions in multiple aspects of intermediary metabolism.

Section III addresses bioenergetics and the role of high-energy phosphates in energy capture and transfer, the oxidation–reduction reactions involved in biologic oxidation, and metabolic details of energy capture via the respiratory chain and oxidative phosphorylation.

Section IV considers the metabolism of carbohydrates via glycolysis, the citric acid cycle, the pentose phosphate pathway, glycogen metabolism, gluconeogenesis, and the control of blood glucose.

Section V outlines the nature of simple and complex lipids, lipid transport and storage, the biosynthesis and degradation of fatty acids and more complex lipids, and the reactions and metabolic regulation of cholesterol biosynthesis and transport in human subjects.

Section VI discusses protein catabolism, urea biosynthesis, and the catabolism of amino acids, and stresses the medically significant metabolic disorders associated with their incomplete catabolism. The final chapter in this section considers the biochemistry of the porphyrins and bile pigments.

Section VII first outlines the structure and function of nucleotides and nucleic acids, and then details DNA replication and repair, RNA synthesis and modification, protein synthesis, the principles of recombinant DNA technology, and the regulation of gene expression.

Section VIII considers aspects of extracellular and intracellular communication. Specific topics include membrane structure and function, the molecular bases of the actions of hormones, and signal transduction.

Sections IX, X, and XI address many topics of significant medical importance.

Section IX discusses nutrition, digestion, and absorption, micronutrients including, vitamins, free radicals and antioxidants, glycoproteins, the metabolism of xenobiotics, and clinical biochemistry.

Section X addresses intracellular traffic and the sorting of proteins, the extracellular matrix, muscle and the cytoskeleton, plasma proteins and immunoglobulins, and the biochemistry of red cells and of white cells.

Section XI includes hemostasis and thrombosis, an overview of cancer, the biochemistry of aging, and a selection of case histories.

Acknowledgments

The authors thank Michael Weitz for his role in the planning of this edition and Peter Boyle for overseeing its preparation for publication. We also thank Tasneem Kauser and her colleagues at KnowledgeWorks Global Ltd. for their efforts in managing editing, typesetting, and artwork. We gratefully acknowledge numerous suggestions and corrections received from students and colleagues from around the world.

Peter J. Kennelly
Kathleen M. Botham
Owen P. McGuinness
Victor W. Rodwell
P. Anthony Weil

SECTION I

Structures & Functions of Proteins & Enzymes

CHAPTER 1

Biochemistry & Medicine

Victor W. Rodwell, PhD

OBJECTIVES

After studying this chapter, you should be able to:

- Understand the importance of the ability of cell-free extracts of yeast to ferment sugars, an observation that enabled discovery of the intermediates of fermentation, glycolysis, and other metabolic pathways.
- Appreciate the scope of biochemistry and its central role in the life sciences, and that biochemistry and medicine are intimately related disciplines.
- Appreciate that biochemistry integrates knowledge of the chemical processes in living cells with strategies to maintain health, understand disease, identify potential therapies, and enhance our understanding of the origins of life on earth.
- Describe how genetic approaches have been critical for elucidating many areas of biochemistry, and how the Human Genome Project has furthered advances in numerous aspects of biology and medicine.

BIOMEDICAL IMPORTANCE

Biochemistry and medicine enjoy a mutually cooperative relationship. Biochemical studies have illuminated many aspects of health and disease, and the study of various aspects of health and disease has opened up new areas of biochemistry. The medical relevance of biochemistry both in normal and abnormal situations is emphasized throughout this book. Biochemistry makes significant contributions to the fields of cell biology, physiology, immunology, microbiology, pharmacology, toxicology, and epidemiology, as well as the fields of inflammation, cell injury, and cancer. These close relationships emphasize that life, as we know it, depends on biochemical reactions and processes.

DISCOVERY THAT A CELL-FREE EXTRACT OF YEAST CAN FERMENT SUGAR

Although the ability of yeast to "ferment" various sugars to ethyl alcohol has been known for millennia, only comparatively recently did this process initiate the science of biochemistry. The great French microbiologist Louis Pasteur maintained that fermentation could only occur in intact cells. However, in 1899, the brothers Büchner discovered that fermentation could occur in the *absence* of intact cells when they stored a yeast extract in a crock of concentrated sugar solution, added as a preservative. Overnight, the contents of the crock fermented, spilled over the laboratory bench and floor,

and dramatically demonstrated that fermentation can proceed in the absence of an intact cell. This discovery unleashed an avalanche of research that initiated the science of biochemistry. Investigations revealed the vital roles of inorganic phosphate, ADP, ATP, and NAD(H), and ultimately identified the phosphorylated sugars and the chemical reactions and enzymes that convert glucose to pyruvate (glycolysis) or to ethanol and CO_2 (fermentation). Research beginning in the 1930s identified the intermediates of the citric acid cycle and of urea biosynthesis, and revealed the essential roles of certain vitamin-derived cofactors or "coenzymes" such as thiamin pyrophosphate, riboflavin, and ultimately coenzyme A, coenzyme Q, and cobamide coenzyme. The 1950s revealed how complex carbohydrates are synthesized from, and broken down into simple sugars, and the pathways for biosynthesis of pentoses, and the catabolism of amino acids and fatty acids.

Investigators employed animal models, perfused intact organs, tissue slices, cell homogenates and their subfractions, and subsequently purified enzymes. Advances were enhanced by the development of analytical ultracentrifugation, paper and other forms of chromatography, and the post-World War II availability of radioisotopes, principally ^{14}C, ^{3}H, and ^{32}P, as "tracers" to identify the intermediates in complex pathways such as that of cholesterol biosynthesis. X-ray crystallography was then used to solve the three-dimensional structures of numerous proteins, polynucleotides, enzymes, and viruses. Genetic advances that followed the realization that DNA was a double helix include the polymerase chain reaction, and transgenic animals or those with gene knockouts. The methods used to prepare, analyze, purify, and identify metabolites and the activities of natural and recombinant enzymes and their three-dimensional structures are discussed in the following chapters.

BIOCHEMISTRY & MEDICINE HAVE PROVIDED MUTUAL ADVANCES

The two major concerns for workers in the health sciences—and particularly physicians—are the understanding and maintenance of health and effective treatment of disease. Biochemistry impacts both of these fundamental concerns, and the interrelationship of biochemistry and medicine is a wide, two-way street. Biochemical studies have illuminated many aspects of health and disease, and conversely, the study of various aspects of health and disease has opened up new areas of biochemistry (**Figure 1–1**). An early example of how investigation of protein structure and function revealed the single difference in amino acid sequence between normal hemoglobin and sickle cell hemoglobin. Subsequent analysis of numerous variant sickle cell and other hemoglobins has contributed significantly to our understanding of the structure and function both of hemoglobin and of other proteins. During the early 1900s, the English physician Archibald Garrod studied patients with the relatively rare disorders of alkaptonuria, albinism, cystinuria, and pentosuria, and established that these conditions were genetically determined. Garrod designated these conditions as **inborn errors of metabolism**. His insights provided a foundation for the development of the field of human biochemical genetics. A more recent example was investigation of the genetic and molecular basis of familial hypercholesterolemia, a disease that results in early-onset atherosclerosis. In addition to clarifying different genetic mutations responsible for this disease, this provided a deeper understanding of cell receptors and mechanisms of uptake, not only of cholesterol but also of how other molecules cross cell membranes. Studies of **oncogenes** and **tumor suppressor genes** in cancer cells have directed attention to the molecular mechanisms involved in the control of normal cell growth. These examples illustrate how the study of disease can open up areas of basic biochemical research. Science provides physicians and other workers in health care and biology with a foundation that impacts practice, stimulates curiosity, and promotes the adoption of scientific approaches for continued learning.

BIOCHEMICAL PROCESSES UNDERLIE HUMAN HEALTH

Biochemical Research Impacts Nutrition & Preventive Medicine

The World Health Organization (WHO) defines health as a state of "complete physical, mental, and social well-being and not merely the absence of disease and infirmity." From a biochemical

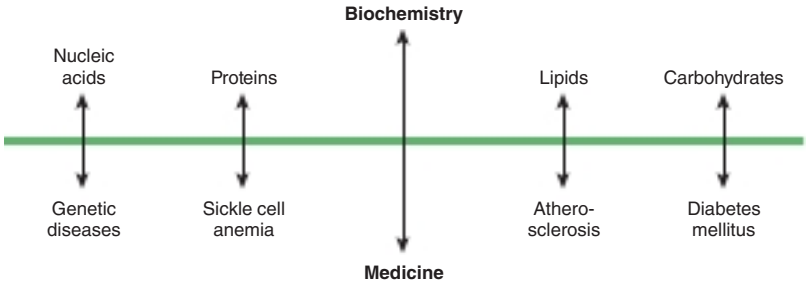


FIGURE 1–1 A two-way street connects biochemistry and medicine. Knowledge of the biochemical topics listed above the green line of the diagram has clarified our understanding of the diseases shown below the green line. Conversely, analyses of the diseases have cast light on many areas of biochemistry. Note that sickle cell anemia is a genetic disease, and that both atherosclerosis and diabetes mellitus have genetic components.

viewpoint, health may be considered that situation in which all of the many thousands of intra- and extracellular reactions that occur in the body are proceeding at rates commensurate with the organism's survival under pressure from both internal and external challenges. The maintenance of health requires optimal dietary intake of **vitamins**, certain **amino acids** and **fatty acids**, various **minerals**, and **water**. Understanding nutrition depends to a great extent on knowledge of biochemistry, and the sciences of biochemistry and nutrition share a focus on these chemicals. Recent increasing emphasis on systematic attempts to maintain health and forestall disease, or **preventive medicine**, includes nutritional approaches to the prevention of diseases such as atherosclerosis and cancer.

Most Diseases Have a Biochemical Basis

Apart from infectious organisms and environmental pollutants, many diseases are manifestations of abnormalities in genes, proteins, chemical reactions, or biochemical processes, each of which can adversely affect one or more critical biochemical functions. Examples of disturbances in human biochemistry responsible for diseases or other debilitating conditions include electrolyte imbalance, defective nutrient ingestion or absorption, hormonal imbalances, toxic chemicals or biologic agents, and DNA-based genetic disorders. To address these challenges, biochemical research continues to be interwoven with studies in disciplines such as genetics, cell biology, immunology, nutrition, pathology, and pharmacology. In addition, many biochemists are vitally interested in contributing to solutions to key issues such as the ultimate survival of mankind, and educating the public to support use of the scientific method in solving environmental and other major problems that confront our civilization.

Impact of the Human Genome Project on Biochemistry, Biology, & Medicine

Initially unanticipated rapid progress in the late 1990s in sequencing the human genome led in the mid-2000s to the announcement that over 90% of the genome had been sequenced. This effort was headed by the International Human Genome Sequencing Consortium and by Celera Genomics. Except for a few gaps, the sequence of the entire human genome was completed in 2003, just 50 years after the description of the double-helical nature of DNA by Watson and Crick. The implications for biochemistry, medicine, and indeed for all of biology, are virtually unlimited. For example, the ability to isolate and sequence a gene and to investigate its structure and function by sequencing and "gene knockout" experiments have revealed previously unknown genes and their products, and new insights have been gained concerning human evolution and procedures for identifying disease-related genes.

Major advances in biochemistry and understanding human health and disease continue to be made by mutation of the genomes of model organisms such as yeast, the fruit fly *Drosophila melanogaster*, the roundworm *Caenorhabditis elegans*, and the zebra fish; all organisms that can be genetically manipulated to provide insight into the functions of individual genes. These advances can potentially provide clues to curing human diseases such as cancer and Alzheimer disease. **Figure 1–2** highlights areas that have developed or accelerated as a direct result of progress made in the Human Genome Project (HGP). New "**-omics**" fields focus on comprehensive study of the structures and functions of the molecules with which each is concerned. The products of genes (RNA molecules and proteins) are being studied using the techniques of **transcriptomics** and **proteomics**. A spectacular example of the speed of progress in transcriptomics is the explosion of knowledge about small RNA molecules as regulators of gene activity. Other -omics fields include **glycomics**, **lipidomics**, **metabolomics**, **nutrigenomics**, and **pharmacogenomics**. To keep pace with the information generated, **bioinformatics** has received much attention. Other related fields to which the impetus from the HGP has carried over are **biotechnology**, **bioengineering**, **biophysics**, and **bioethics**.

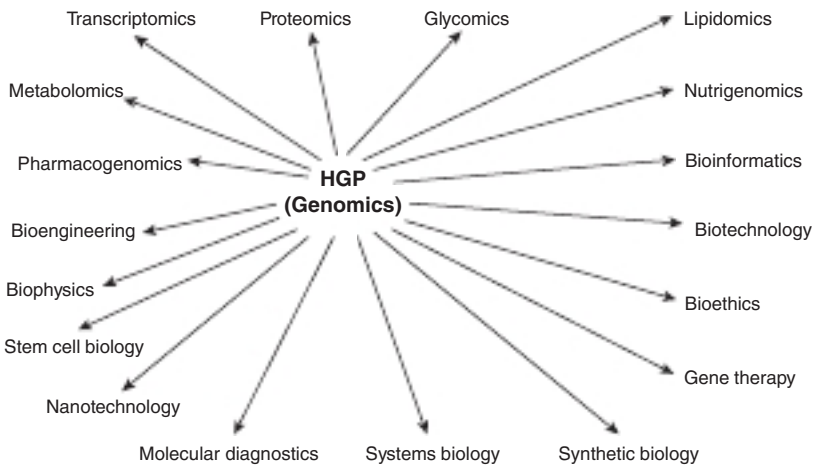


FIGURE 1–2 The Human Genome Project (HGP) has influenced many disciplines and areas of research. Biochemistry is not listed since it predates commencement of the HGP, but disciplines such as bioinformatics, genomics, glycomics, lipidomics, metabolomics, molecular diagnostics, proteomics, and transcriptomics are nevertheless active areas of biochemical research.

Definitions of these -omics fields and other terms appear in the Glossary of this chapter. **Nanotechnology** is an active area, which, for example, may provide novel methods of diagnosis and treatment for cancer and other disorders. **Stem cell biology** is at the center of much current research. **Gene therapy** has yet to deliver the promise that it appears to offer, but it seems probable that ultimately will occur. Many new **molecular diagnostic tests** have developed in areas such as genetic, microbiologic, and immunologic testing and diagnosis. **Systems biology** is also burgeoning. The outcomes of research in the various areas mentioned above will impact tremendously the future of biology, medicine, and the health sciences. **Synthetic biology** offers the potential for creating living organisms, initially small bacteria, from genetic material in vitro that might carry out specific tasks such as cleansing petroleum spills. All of the above make the 21st century an exhilarating time to be directly involved in biology and medicine.

SUMMARY

- Biochemistry is the science concerned with the molecules present in living organisms, individual chemical reactions and their enzyme catalysts, and the expression and regulation of each metabolic process. Biochemistry has become the basic language of all biologic sciences.
- Despite the focus on human biochemistry in this text, biochemistry concerns the entire spectrum of life forms, from viruses, bacteria, and plants to complex eukaryotes such as human beings.
- Biochemistry, medicine, and other health care disciplines are intimately related. Health in all species depends on a harmonious balance of the biochemical reactions occurring in the body, while disease reflects abnormalities in biomolecules, biochemical reactions, or biochemical processes.
- Advances in biochemical knowledge have illuminated many areas of medicine, and the study of diseases has often revealed previously unsuspected aspects of biochemistry.
- Biochemical approaches are often fundamental in illuminating the causes of diseases and in designing appropriate therapy. Biochemical laboratory tests also represent an integral component of diagnosis and monitoring of treatment.
- A sound knowledge of biochemistry and of other related basic disciplines is essential for the rational practice of medicine and related health sciences.
- Results of the HGP and of research in related areas will have a profound influence on the future of biology, medicine, and other health sciences.
- Genomic research on model organisms such as yeast, the fruit fly *D. melanogaster*, the roundworm *C. elegans*, and the zebra fish provides insight into understanding human diseases.

GLOSSARY

Bioengineering: The application of engineering to biology and medicine.

Bioethics: The area of ethics that is concerned with the application of moral and ethical principles to biology and medicine.

Bioinformatics: The discipline concerned with the collection, storage, and analysis of biologic data, for example, DNA, RNA, and protein sequences.

Biophysics: The application of physics and its techniques to biology and medicine.

Biotechnology: The field in which biochemical, engineering, and other approaches are combined to develop biologic products of use in medicine and industry.

Gene Therapy: Applies to the use of genetically engineered genes to treat various diseases.

Genomics: The genome is the complete set of genes of an organism, and genomics is the in-depth study of the structures and functions of genomes.

Glycomics: The glycome is the total complement of simple and complex carbohydrates in an organism. Glycomics is the systematic study of the structures and functions of glycomes such as the human glycome.

Lipidomics: The lipidome is the complete complement of lipids found in an organism. Lipidomics is the in-depth study of the structures and functions of all members of the lipidome and their interactions, in both health and disease.

Metabolomics: The metabolome is the complete complement of metabolites (small molecules involved in metabolism) present in an organism. Metabolomics is the in-depth study of their structures, functions, and changes in various metabolic states.

Molecular Diagnostics: Refers to the use of molecular approaches such as DNA probes to assist in the diagnosis of various biochemical, genetic, immunologic, microbiologic, and other medical conditions.

Nanotechnology: The development and application to medicine and to other areas of devices such as nanoshells, which are only a few nanometers in size (10^{-9} m = 1 nm).

Nutrigenomics: The systematic study of the effects of nutrients on genetic expression and of the effects of genetic variations on the metabolism of nutrients.

Pharmacogenomics: The use of genomic information and technologies to optimize the discovery and development of new drugs and drug targets.

Proteomics: The proteome is the complete complement of proteins of an organism. Proteomics is the systematic study of the structures and functions of proteomes and their variations in health and disease.

Stem Cell Biology: Stem cells are undifferentiated cells that have the potential to self-renew and to differentiate into any of the adult cells of an organism. Stem cell biology concerns the biology of stem cells and their potential for treating various diseases.

Synthetic Biology: The field that combines biomolecular techniques with engineering approaches to build new biologic functions and systems.

Systems Biology: The field concerns complex biologic systems studied as integrated entities.

Transcriptomics: The comprehensive study of the transcriptome, the complete set of RNA transcripts produced by the genome during a fixed period of time.

APPENDIX

Shown are selected examples of databases that assemble, annotate, and analyze data of biomedical importance.

ENCODE: ENCyclopedia Of DNA Elements. A collaborative effort that combines laboratory and computational approaches to identify every functional element in the human genome.

GenBank: Protein sequence database of the National Institutes of Health (NIH) stores all known biologic nucleotide sequences and their translations in a searchable form.

HapMap: Haplotype **Map**, an international effort to identify single nucleotide polymorphisms (SNPs) associated with common human diseases and differential responses to pharmaceuticals.

ISDB: International **S**equence **D**ata**B**ase that incorporates DNA databases of Japan and of the European Molecular Biology Laboratory (EMBL).

PDB: Protein **D**ata**B**ase. Three-dimensional structures of proteins, polynucleotides, and other macromolecules, including proteins bound to substrates, inhibitors, or other proteins.

CHAPTER

2

Water & pH

Peter J. Kennelly, PhD, & Victor W. Rodwell, PhD

OBJECTIVES

After studying this chapter, you should be able to:

- Describe the properties of water that account for its surface tension, viscosity, liquid state at ambient temperature, and solvent power.
- Represent the structures of organic compounds that can serve as hydrogen bond donors or acceptors.
- Explain the role played by entropy in the association and orientation, in an aqueous environment, of hydrophobic and amphipathic molecules.
- Indicate the quantitative contributions of salt bridges, hydrophobic interactions, and van der Waals forces to stabilizing the 3-D conformation of macromolecules.
- Explain the relationship of pH to acidity, alkalinity, and the quantitative determinants that characterize weak and strong acids.
- Calculate the shift in pH that accompanies the addition of a given quantity of acid or base to a buffered solution.
- Describe what buffers do, how they do it, and the conditions under which a buffer is most effective under physiologic or other conditions.
- Use the Henderson-Hasselbalch equation to calculate the net charge on a polyelectrolyte at a given pH.

BIOMEDICAL IMPORTANCE

Water is the predominant chemical component of living organisms. Its unique physical properties, which include the ability to solvate a wide range of organic and inorganic molecules, derive from water's dipolar structure and exceptional capacity for forming hydrogen bonds. The manner in which water interacts with a solvated biomolecule influences the structure both of the biomolecule and of water itself. An excellent nucleophile, water is a reactant or product in many metabolic reactions. Regulation of water balance depends on hypothalamic mechanisms that control thirst, on antidiuretic hormone (ADH), on retention or excretion of water by the kidneys, and on evaporative loss. Nephrogenic diabetes insipidus, which involves the inability to concentrate urine or adjust to subtle changes in extracellular fluid osmolarity, results from the unresponsiveness of renal tubular osmoreceptors to ADH.

Water has a slight propensity to dissociate into hydroxide ions and protons. The concentration of protons, or **acidity**, of aqueous solutions is generally reported using the logarithmic pH scale. Bicarbonate and other buffers normally maintain the pH of extracellular fluid between 7.35 and 7.45. Suspected disturbances of acid-base balance are verified by measuring the pH of arterial blood and the CO_2 content of venous blood. Causes of acidosis (blood pH <7.35) include diabetic ketosis and lactic acidosis. Alkalosis (pH >7.45) may follow vomiting of acidic gastric contents.

WATER IS AN IDEAL BIOLOGIC SOLVENT

Water Molecules Form Dipoles

A water molecule is an irregular, slightly skewed tetrahedron with oxygen at its center (**Figure 2–1**). The corners are occupied by the two hydrogens and the unshared electrons of the remaining two sp^3-hybridized orbitals of oxygen. The 105° angle between the two hydrogen atoms differs slightly from the ideal tetrahedral angle, 109.5°. The strongly electronegative oxygen atom in a water molecule attracts electrons away from the hydrogen nuclei, leaving them with a partial

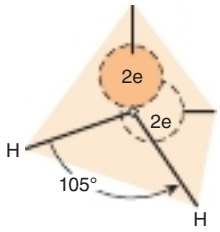


FIGURE 2–1 The water molecule has tetrahedral geometry.

positive charge, while its two unshared electron pairs constitute a region of local negative charge. This asymmetric charge distribution is referred to as a **dipole**.

Water's strong dipole is responsible for its high **dielectric constant**. As described quantitatively by Coulomb's law, the strength of interaction F between oppositely charged particles is inversely proportionate to the dielectric constant ε of the surrounding medium. The dielectric constant for a vacuum is essentially unity; for hexane it is 1.9; for ethanol, 24.3; and for water at 25°C, 78.5. When dissolved in water, the force of attraction between charged and polar species is greatly decreased relative to solvents with lower dielectric constants. Its strong dipole and high dielectric constant enable water to dissolve large quantities of charged compounds such as salts.

Water Molecules Form Hydrogen Bonds

A partially unshielded hydrogen nucleus covalently bound to an electron-withdrawing oxygen or nitrogen atom can interact with an unshared electron pair on another oxygen or nitrogen atom to form a **hydrogen bond**. Since water molecules contain both of these features, hydrogen bonding favors the self-association of water molecules into ordered arrays (**Figure 2–2**). On average, each molecule in liquid water associates through hydrogen bonds with 3.5 others. These bonds are both relatively weak and transient, with a half-life of a few picoseconds. Rupture of a hydrogen bond in liquid water requires only about 4.5 kcal/mol, less than 5% of the energy required to rupture a covalent O—H bond. The exceptional capacity of this relatively small, 18 g/mol, molecule to form hydrogen bonds profoundly influences the physical properties of water and accounts for its high viscosity, surface tension, and boiling point.

Hydrogen bonding enables water to dissolve many organic biomolecules that contain functional groups which

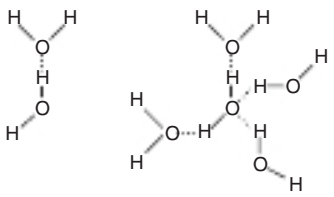


FIGURE 2–2 Water molecules self-associate via hydrogen bonds. Shown are the association of two water molecules (**left**) and a hydrogen-bonded cluster of four water molecules (**right**). Notice that water can serve simultaneously both as a hydrogen donor and as a hydrogen acceptor.

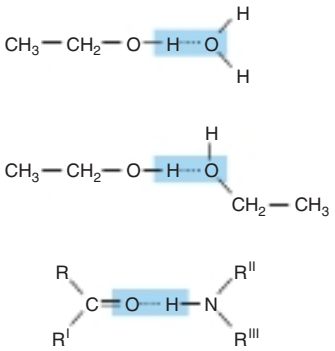


FIGURE 2–3 Additional polar groups participate in hydrogen bonding. Shown are hydrogen bonds formed between alcohol and water, between two molecules of ethanol, and between the peptide carbonyl oxygen and the peptide nitrogen hydrogen of an adjacent amino acid.

can participate in hydrogen bonding. The oxygen atoms of aldehydes, ketones, and amides, for example, provide lone pairs of electrons that can serve as hydrogen acceptors. Alcohols, carboxylic acids, and amines can serve both as hydrogen acceptors and as donors of unshielded hydrogen atoms for formation of hydrogen bonds (**Figure 2–3**).

INTERACTION WITH WATER INFLUENCES THE STRUCTURE OF BIOMOLECULES

Covalent & Noncovalent Bonds Stabilize Biologic Molecules

The covalent bond is the strongest force that holds molecules together (**Table 2–1**). Noncovalent forces, while of lesser magnitude, predominate in stabilizing the folding of the polypeptides and other macromolecules into the complex three-dimensional conformations essential to their functional competence (see Chapter 5) as well as the association of biomolecules into multicomponent complexes. Examples of the latter include the coalescence of the polypeptide subunits that form the hemoglobin tetramer (see Chapter 6); the association

TABLE 2–1 Bond Energies for Atoms of Biologic Significance

Bond Type	Energy (kcal/mol)	Bond Type	Energy (kcal/mol)
O—O	34	O=O	96
S—S	51	C—H	99
C—N	70	C=S	108
S—H	81	O—H	110
C—C	82	C=C	147
C—O	84	C=N	147
N—H	94	C=O	164

of the two polynucleotide strands that comprise a DNA double helix (see Chapter 34); and the coalescence of billions of phospholipid, glycosphingolipid, cholesterol, and other molecules into the bilayer that constitutes the foundation of the plasma membrane of an animal cell (see Chapter 40). These forces, which can be either attractive or repulsive, involve interactions both within the biomolecule and, most importantly, between it and the water that forms the principal component of the surrounding environment.

In Water, Biomolecules Fold to Position Hydrophobic Groups Within Their Interior

Most biomolecules are **amphipathic**; that is, they possess regions rich in charged or polar functional groups as well as regions with hydrophobic character. Proteins tend to fold with the R-groups of amino acids with hydrophobic side chains in the interior. Amino acids with charged or polar amino acid side chains (eg, arginine, glutamate, serine; see Table 3–1) generally are present on the surface in contact with water. A similar pattern prevails in a phospholipid bilayer where the charged "head groups" of phosphatidylserine or phosphatidylethanolamine contact water while their hydrophobic fatty acyl side chains cluster together, excluding water (see Figure 40–5). This pattern minimizes energetically unfavorable contacts between water and hydrophobic groups. It also maximizes the opportunities for the formation of energetically favorable charge-dipole, dipole-dipole, and hydrogen bonding interactions between polar groups on the biomolecule and water.

Hydrophobic Interactions

Hydrophobic interaction refers to the tendency of nonpolar compounds to self-associate in an aqueous environment. This self-association is driven neither by mutual attraction nor by what are sometimes incorrectly referred to as "hydrophobic bonds." Self-association minimizes the disruption of energetically favorable interactions between and is therefore driven by the surrounding water molecules.

While the hydrogen atoms of nonpolar groups such as the methylene groups of hydrocarbons do not form hydrogen bonds, they do affect the structure of the water with which they are in contact. Water molecules adjacent to a hydrophobic group are restricted in the number of orientations (degrees of freedom) that permit them to participate in the maximum number of energetically favorable hydrogen bonds. Maximal formation of multiple hydrogen bonds, which maximizes enthalpy, can be maintained only by increasing the order of the adjacent water molecules, with an accompanying decrease in entropy.

It follows from the second law of thermodynamics that the optimal free energy of a hydrocarbon-water mixture is a function of both maximal enthalpy (from hydrogen bonding) and highest entropy (maximum degrees of freedom). Thus, nonpolar molecules tend to form droplets that minimize exposed surface area and reduce the number of water molecules whose motional freedom becomes restricted (**Figure 2–4**). Similarly, in the aqueous environment of the living cell the hydrophobic portions of amphipathic biopolymers tend to be buried inside the structure of the molecule, or within a lipid bilayer, minimizing contact with water.

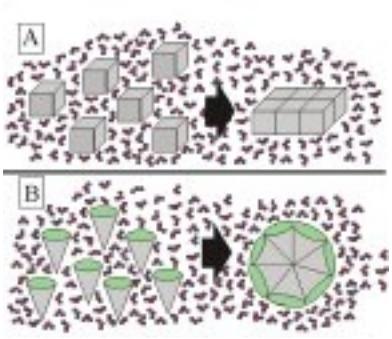


FIGURE 2–4 Hydrophobic interactions are driven by the surrounding water molecules. Water molecules are represented by one red (oxygen) and two blue (hydrogen) circles. The hydrophobic surfaces of solute molecules are colored gray and, where present, hydrophilic ones are colored green. **A.** When the six hydrophobic cubes shown are dispersed in water (**left**), the surrounding water molecules (red oxygens and blue hydrogens) are forced to engage in entropically unfavorable interactions with all 36 faces of the cubes. However, when the six hydrophobic cubes aggregate together (**right**), the number of exposed faces is reduced to 22. The aggregate forms and its stability is maintained, not by some attractive force, but because aggregation reduces the number of water molecules that are unfavorably affected by nearly 40%. **B.** Amphipathic molecules associate together for the same reason. However, the structure of the resulting complex (eg, micelle or bilayer) is determined by the geometries of the hydrophobic (gray) and hydrophilic (green) regions.

Electrostatic Interactions

Electrostatic interactions between oppositely charged groups within or between biomolecules are termed **salt bridges**. Salt bridges are comparable in strength to hydrogen bonds but act over larger distances. They therefore often facilitate the binding of charged molecules and ions to proteins and nucleic acids.

van der Waals Forces

van der Waals forces arise from attractions between transient dipoles generated by the rapid movement of electrons in all neutral atoms. Significantly weaker than hydrogen bonds but potentially extremely numerous, van der Waals forces decrease as the sixth power of the distance separating atoms (**Figure 2–5**). Thus, they act over very short distances, typically 2 to 4 Å.

Multiple Forces Stabilize Biomolecules

The DNA double helix illustrates the contribution of multiple forces to the structure of biomolecules. While each individual DNA strand is held together by covalent bonds, the two strands of the helix are held together exclusively by noncovalent

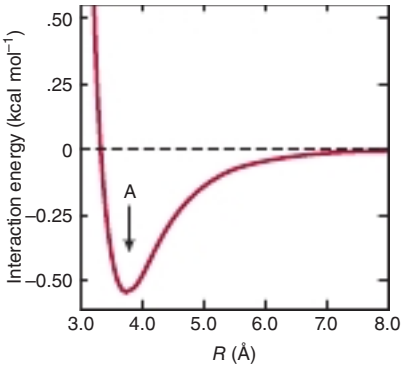


FIGURE 2–5 The strength of van der Waals interactions varies with the distance, *R*, between interacting species. The force of interaction between interacting species increases with decreasing distance between them until they are separated by the van der Waals contact distance (see arrow marked A). Repulsion due to interaction between the electron clouds of each atom or molecule then supervenes. While individual van der Waals interactions are extremely weak, their cumulative effect is nevertheless substantial for macromolecules such as DNA and proteins which have many atoms in close contact.

interactions such as hydrogen bonds between nucleotide bases (Watson-Crick base pairing) and van der Waals interactions between the stacked purine and pyrimidine bases. The double helix presents the charged phosphate groups and polar hydroxyl groups from the ribose sugars of the DNA backbone to water while burying the relatively hydrophobic nucleotide bases inside. The extended backbone maximizes the distance between negatively charged phosphates, minimizing unfavorable electrostatic interactions (see Figure 34–2).

WATER IS AN EXCELLENT NUCLEOPHILE

Metabolic reactions often involve the attack by lone pairs of electrons residing on electron-rich molecules termed **nucleophiles** upon electron-poor atoms called **electrophiles**. Nucleophiles and electrophiles do not necessarily possess a formal negative or positive charge. Water, whose two lone pairs of sp^3 electrons bear a partial negative charge (see Figure 2–1), is an excellent nucleophile. Other nucleophiles of biologic importance include the oxygen atoms of phosphates, alcohols, and carboxylic acids; the sulfur of thiols; and the nitrogen atoms of amines and of the imidazole ring of histidine. Common electrophiles include the carbonyl carbons in amides, esters, aldehydes, and ketones and the phosphorus atoms of phosphoesters.

Nucleophilic attack by water typically results in the cleavage of the amide, glycoside, or ester bonds that hold biopolymers together. This process is termed **hydrolysis**. Conversely, when monomer units such as amino acids or monosaccharides are joined or **condensed** together to form biopolymers, such as proteins or starch, water is a product.

Hydrolysis typically is a thermodynamically favored reaction. Yet, the amide and phosphoester bonds of polypeptides and oligonucleotides are stable in the aqueous environment of the cell. This seemingly paradoxical behavior reflects the fact that the thermodynamics that govern the equilibrium point of a reaction do not determine the *rate* at which it will proceed toward its equilibrium point. In the cell, macromolecular catalysts called **enzymes** accelerate the rate of hydrolytic and other chemical reactions when needed. **Proteases** catalyze the hydrolysis of proteins into their component amino acids, while **nucleases** catalyze the hydrolysis of the phosphoester bonds in DNA and RNA. Precise and differential control of enzyme activity, including the sequestration of enzymes in specific organelles, enables cells to determine the physiologic circumstances under which a given biopolymer will be synthesized or degraded.

Many Metabolic Reactions Involve Group Transfer

Many of the enzymic reactions responsible for synthesis and breakdown of biomolecules involve the transfer of a chemical group G from a donor D to an acceptor A to form an acceptor group complex, A—G:

$$D—G + A \rightleftarrows A—G + D$$

The hydrolysis and phosphorolysis of glycogen, for example, involve the transfer of glucosyl groups to water or to orthophosphate. Since the equilibrium constants for these hydrolysis reactions strongly favor the formation of split products, it follows that many of the group transfer reactions responsible for the biosynthesis of macromolecules are, in and of themselves, thermodynamically unfavored. Enzyme catalysts play a critical role in surmounting these barriers by virtue of their capacity to directly link two normally separate reactions together. For example, by linking an energetically unfavorable group transfer reaction to a thermodynamically favorable one such as the hydrolysis of ATP, a new enzyme-catalyzed reaction can be generated. The free energy change of this coupled reaction will be the sum of the individual values for the two that were linked, one whose net *overall* change in free energy favors the formation of the covalent bonds required for biopolymer synthesis.

Water Molecules Exhibit a Slight but Important Tendency to Dissociate

The ability of water to ionize, while slight, is of central importance for life. Since water can act both as an acid and as a base, its ionization may be represented as an intermolecular proton transfer that forms a hydronium ion (H_3O^+) and a hydroxide ion (OH^-):

$$H_2O + H_2O \rightleftarrows H_3O + OH^-$$

The transferred proton is actually associated with a cluster of water molecules. Protons exist in solution not only as H_3O^+ but also as multimers such as $H_5O_2^+$ and $H_7O_3^+$. The proton is nevertheless routinely represented as H^+, even though it is in fact highly hydrated.

Since hydronium and hydroxide ions continuously recombine to form water molecules, an *individual* hydrogen or oxygen cannot be stated to be present as an ion or as part of a water molecule. At one instant it is an ion; an instant later it is part of a water molecule. Individual ions or molecules are therefore not considered. We refer instead to the *probability* that at any instant in time, a given hydrogen will be present as an ion or as part of a water molecule. Since 1 g of water contains 3.35×10^{22} molecules, the ionization of water can be described statistically. To state that the probability that a hydrogen exists as an ion is 0.01 means that at any given moment in time, a hydrogen atom has 1 chance in 100 of being an ion and 99 chances out of 100 of being part of a water molecule. The actual probability of a hydrogen atom in pure water existing as a hydrogen ion is approximately 1.8×10^{-9}. The probability of its being part of a water molecule thus is almost unity. Stated another way, for every hydrogen ion or hydroxide ion in pure water, there are 0.56 billion or 0.56×10^9 water molecules. Hydrogen ions and hydroxide ions nevertheless contribute significantly to the properties of water.

For dissociation of water,

$$K = \frac{[H^+][OH^-]}{[H_2O]}$$

where the brackets represent molar concentrations (strictly speaking, molar activities) and K is the **dissociation constant**. Since 1 mole (mol) of water weighs 18 g, 1 liter (L) (1000 g) of water contains $1000 \div 18 = 55.56$ mol. Pure water thus is 55.56 molar. Since the probability that a hydrogen in pure water will exist as a hydrogen ion is 1.8×10^{-9}, the molar concentration of H^+ ions (or of OH^- ions) in pure water is the product of the probability, 1.8×10^{-9}, times the molar concentration of water, 55.56 mol/L. The result is 1.0×10^{-7} mol/L.

We can now calculate the dissociation constant K for pure water:

$$K = \frac{[H^+][OH^-]}{[H_2O]} = \frac{[10^{-7}][10^{-7}]}{[55.56]}$$
$$= 0.018 \times 10^{-14} = 1.8 \times 10^{-16}\ \text{mol/L}$$

The molar concentration of water, 55.56 mol/L, is too great to be significantly affected by dissociation. It is therefore considered to be essentially constant. The concentration of pure water may therefore be incorporated into the dissociation constant K to provide a useful new constant K_w termed the **ion product** for water:

$$K = \frac{[H^+][OH^-]}{[H_2O]} = 1.8 \times 10^{-16}\ \text{mol/L}$$
$$K_w = (K)[H_2O] = [H^+][OH^-]$$
$$= (1.8 \times 10^{-16}\ \text{mol/L})(55.56\ \text{mol/L})$$
$$= 1.00 \times 10^{-14}\ (\text{mol/L})^2$$

Note that the dimensions of K are moles per liter and those of K_w are moles^2 per liter^2. As its name suggests, the ion product K_w is numerically equal to the product of the molar concentrations of H^+ and OH^-:

$$K_w = [H^+][OH^-]$$

At 25°C, $K_w = (10^{-7})^2$, or 10^{-14} (mol/L)^2. At temperatures below 25°C, K_w is somewhat less than 10^{-14}, and at temperatures above 25°C it is somewhat greater than 10^{-14}. Within the stated limitations of temperature, K_w equals 10^{-14} (mol/L)^2 for all aqueous solutions, even solutions containing acids or bases. We can therefore use K_w to calculate the pH of any aqueous solution.

pH IS THE NEGATIVE LOG OF THE HYDROGEN ION CONCENTRATION

The term **pH** was introduced in 1909 by Sörensen, who defined it as the negative log of the hydrogen ion concentration:

$$pH = -\log[H^+]$$

This definition, while not rigorous, suffices for most biochemical purposes. To calculate the pH of a solution:

1. Calculate the hydrogen ion concentration $[H^+]$.
2. Calculate the base 10 logarithm of $[H^+]$.
3. pH is the negative of the value found in step 2.

For example, for pure water at 25°C,

$$pH = -\log[H^+] = -\log 10^{-7} = -(-7) = 7.0$$

This value is also known as the *power* (English), *puissant* (French), or *potennz* (German) of the exponent, hence the use of the term "p."

Low pH values correspond to high concentrations of H^+ and high pH values correspond to low concentrations of H^+.

Acids are **proton donors** and bases are **proton acceptors**. **Strong acids** (eg, HCl, H_2SO_4) completely dissociate into anions and protons even in strongly acidic solutions (low pH). **Weak acids** dissociate only partially in acidic solutions. Similarly, **strong bases** (eg, KOH, NaOH), but not **weak bases** like $Ca(OH)_2$, are completely dissociated even at high pH. Many biochemicals are weak acids. Exceptions include phosphorylated intermediates, whose phosphoryl group contains two dissociable protons, the first of which is strongly acidic.

The following examples illustrate how to calculate the pH of acidic and basic solutions.

Example 1: What is the pH of a solution whose hydrogen ion concentration is 3.2×10^{-4} mol/L?

$$pH = -\log[H^+]$$
$$= -\log(3.2 \times 10^{-4})$$
$$= -\log(3.2) - \log(10^{-4})$$
$$= -0.5 + 4.0$$
$$= 3.5$$

Example 2: What is the pH of a solution whose hydroxide ion concentration is 4.0×10^{-4} mol/L? We first define a quantity **pOH** that is equal to $-\log[OH^-]$ and that may be derived from the definition of K_w:

$$K_w = [H^+][OH^-] = 10^{-14}$$

Therefore,

$$\log[H^+] + \log[OH^-] = \log 10^{-14}$$

or

$$pH + pOH = 14$$

To solve the problem by this approach:

$$\begin{aligned}[OH^-] &= 4.0\times10^{-4}\\ pOH &= -\log[OH^-]\\ &= -\log(4.0\times10^{-4})\\ &= -\log(4.0) - \log(10^{-4})\\ &= -0.60 + 4.0\\ &= 3.4\end{aligned}$$

Now

$$\begin{aligned}pH &= 14 - pOH = 14 - 3.4\\ &= 10.6\end{aligned}$$

Examples 1 and 2 illustrate how the logarithmic pH scale facilitates recording and comparing hydrogen ion concentrations that differ by orders of magnitude from one another, 0.00032 M (pH 3.5) and 0.000000000025 M (pH 10.6).

Example 3: What are the pH values of (a) 2.0×10^{-2} mol/L KOH and of (b) 2.0×10^{-6} mol/L KOH? The OH^- arises from two sources, KOH and water. Since pH is determined by the total $[H^+]$ (and pOH by the total $[OH^-]$), both sources must be considered. In the first case (a), the contribution of water to the total $[OH^-]$ is negligible. The same cannot be said for the second case (b):

	Concentration (mol/L)	
	(a)	(b)
Molarity of KOH	2.0×10^{-2}	2.0×10^{-6}
$[OH^-]$ from KOH	2.0×10^{-2}	2.0×10^{-6}
$[OH^-]$ from water	1.0×10^{-7}	1.0×10^{-7}
Total $[OH^-]$	2.00001×10^{-2}	2.1×10^{-6}

Once a decision has been reached about the significance of the contribution of water, pH may be calculated as shown in Example 3.

The above examples assume that the strong base KOH is completely dissociated in solution and that the concentration of OH^- ions was thus equal to that due to the KOH plus that present initially in the water. This assumption is valid for dilute solutions of strong bases or acids, but not for weak bases or acids. Since weak electrolytes dissociate only slightly in solution, we must use the **dissociation constant** to calculate the concentration of $[H^+]$ (or $[OH^-]$) produced by a given molarity of a weak acid (or base) before calculating total $[H^+]$ (or total $[OH^-]$) and subsequently pH.

Functional Groups That Are Weak Acids Have Great Physiologic Significance

Many biomolecules contain functional groups that are weak acids or bases. Carboxyl groups, amino groups, and phosphate esters, whose second dissociation falls within the physiologic range, are present in proteins and nucleic acids, most coenzymes, and most intermediary metabolites. Knowledge of the dissociation of weak acids and bases thus is basic to understanding the influence of intracellular pH on structure and biologic activity. Charge-based separations such as electrophoresis and ion exchange chromatography are also best understood in terms of the dissociation behavior of functional groups.

When discussing weak acids, we often refer to the protonated species (HA or R—SH) as the **acid** and the unprotonated species (A^- or $R—S^-$) as its **conjugate base**. Similarly, we may refer to the deprotonated form as the **base** (A^- or $R—COO^-$) and the protonated form as its **conjugate acid** (HA or R—COOH).

We express the relative strengths of weak acids in terms of the dissociation constants of the protonated form. Following are the expressions for the dissociation constant (K_a) for a representative weak acid, R—COOH, as well as the conjugate acid, $R—NH_3^+$, of the weak base $R—NH_2$.

$$R—COOH \rightleftarrows R—COO^- + H^+$$

$$K_a = \frac{[R—COO^-][H^+]}{[R—COOH]}$$

$$R—NH_3^+ \rightleftarrows R—NH_2 + H^+$$

$$K_a = \frac{[R—NH_2][H^+]}{[R—NH_3^+]}$$

Since the numeric values of K_a for weak acids are negative exponential numbers, we express K_a as pK_a, where

$$pK_a = -\log K_a$$

Note that pK_a is related to K_a as pH is to $[H^+]$. The stronger the acid, the lower is its pK_a value.

Representative weak acids (left), their conjugate bases (center), and pK_a values (right) include the following:

$R—CH_2—COOH$	$R—CH_2COO^-$	$pK_a = 4-5$
$R—CH_2—NH_3^+$	$R—CH_2—NH_2$	$pK_a = 9-10$
H_2CO_3	HCO_3^-	$pK_a = 6.4$
$H_2PO_4^-$	HPO_4^{-2}	$pK_a = 7.2$

pK_a is used to express the relative strengths of both weak acids and weak bases using a single, unified scale. Under this convention, **the relative strengths of bases are expressed in terms of the pK_a of their conjugate acids**. For polyprotic compounds containing more than one dissociable proton, a numerical subscript is assigned to each dissociation, numbered starting from unity in decreasing order of relative acidity. For a dissociation of the type

$$R—NH_3^+ \rightarrow R—NH_2 + H^+$$

the pK_a is the pH at which the concentration of the acid $R—NH_3^+$ equals that of the base $R—NH_2$.

From the above equations that relate K_a to $[H^+]$ and to the concentrations of undissociated acid and its conjugate base, when

$$[R—COO^-] = [R—COOH]$$

or when

$$[R—NH_2] = [R—NH_3^+]$$

then

$$K_a = [H^+]$$

Thus, when the associated (protonated) and dissociated (conjugate base) species are present at equal concentrations, the prevailing hydrogen ion concentration $[H^+]$ is numerically equal to the dissociation constant, K_a. If the logarithms of both sides of the above equation are taken and both sides are multiplied by −1, the expressions would be as follows:

$$K_a = [H^+]$$
$$-\log K_a = -\log[H^+]$$

Since $-\log K_a$ is defined as pK_a, and $-\log [H^+]$ defines pH, the equation may be rewritten as

$$pK_a = pH$$

that is, **the pK_a of an acid group is the pH at which the protonated and unprotonated species are present at equal concentrations**. The pK_a for an acid may be determined by adding 0.5 equivalent of alkali per equivalent of acid. The resulting pH will equal the pK_a of the acid.

The Henderson-Hasselbalch Equation Describes the Behavior of Weak Acids & Buffers

The Henderson-Hasselbalch equation is derived below.

A weak acid, HA, ionizes as follows:

$$HA \rightleftarrows H^+ + A^-$$

The equilibrium constant for this dissociation is

$$K_a = \frac{[H^+][A^-]}{[HA]}$$

Cross-multiplication gives

$$[H^+][A^-] = K_a[HA]$$

Divide both sides by $[A^-]$:

$$[H^+] = K_a \frac{[HA]}{[A^-]}$$

Take the log of both sides:

$$\log[H^+] = \log\left(K_a \frac{[HA]}{[A^-]}\right)$$
$$= \log K_a + \log\frac{[HA]}{[A^-]}$$

Multiply through by −1:

$$-\log[H^+] = -\log K_a - \log\frac{[HA]}{[A^-]}$$

Substitute pH and pK_a for $-\log [H^+]$ and $-\log K_a$, respectively; then

$$pH = pK_a - \log\frac{[HA]}{[A^-]}$$

Inversion of the last term removes the minus sign and gives the **Henderson-Hasselbalch equation**

$$\mathbf{pH = pK_a + \log\frac{[A^-]}{[HA]}}$$

The Henderson-Hasselbalch equation has great predictive value in protonic equilibria. For example,

1. When an acid is exactly half-neutralized, $[A^-] = [HA]$. Under these conditions,

$$pH = pK_a + \log\frac{[A^-]}{[HA]} = pK_a + \log\left(\frac{1}{1}\right) = pK_a + 0$$

Therefore, at half-neutralization, $pH = pK_a$.

2. When the ratio $[A^-]/[HA] = 100:1$,

$$pH = pK_a + \log\frac{[A^-]}{[HA]}$$
$$pH = pK_a + \log(100/1) = pK_a + 2$$

3. When the ratio $[A^-]/[HA] = 1:10$,

$$pH = pK_a + \log(1/10) = pK_a + (-1)$$

If the equation is evaluated at ratios of $[A^-]/[HA]$ ranging from 10^3 to 10^{-3} and the calculated pH values are plotted, the resulting graph describes the titration curve for a weak acid (**Figure 2–6**).

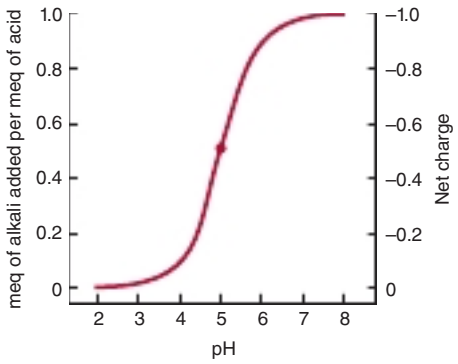


FIGURE 2–6 Titration curve for an acid of the type HA. The heavy dot in the center of the curve indicates the pK_a, 5.0.

Weak Acids Can Be Used to Establish & Maintain the pH of an Aqueous Solution

Solutions of weak acids or bases and their conjugates exhibit **buffering**, the ability to resist a change in pH following addition of strong acid or base. Many metabolic reactions are accompanied by the release or uptake of protons. Oxidative metabolism produces CO_2, the anhydride of carbonic acid, which if not buffered would produce severe acidosis. Biologic maintenance of a constant pH involves buffering by phosphate, bicarbonate, and proteins, which accept or release protons to resist a change in pH. For laboratory experiments using tissue extracts or enzymes, constant pH is maintained by the addition of buffers such as MES ([2-*N*-morpholino]-ethanesulfonic acid, pK_a 6.1), inorganic orthophosphate (pK_{a2} 7.2), HEPES (*N*-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid, pK_a 6.8), or Tris (tris[hydroxymethyl]aminomethane, pK_a 8.3). The value of pK_a relative to the desired pH is the major determinant of which buffer is selected.

Buffering can be observed by using a pH meter while titrating a weak acid or base (see Figure 2–6). We can also calculate the pH shift that accompanies addition of acid or base to a buffered solution. In the following example, the buffered solution (a weak acid, pK_a = 5.0, and its conjugate base) is initially at one of four pH values. We will calculate the pH shift that results when 0.1 meq of KOH is added to 1 meq of each solution:

Initial pH	5.00	5.37	5.60	5.86
$[A^-]_{initial}$	0.50	0.70	0.80	0.88
$[HA]_{initial}$	0.50	0.30	0.20	0.12
$([A^-]/[HA])_{initial}$	1.00	2.33	4.00	7.33
Addition of 0.1 meq of KOH Produces				
$[A^-]_{final}$	0.60	0.80	0.90	0.98
$[HA]_{final}$	0.40	0.20	0.10	0.02
$([A^-]/[HA])_{final}$	1.50	4.00	9.00	49.0
log $([A^-]/[HA])_{final}$	0.18	0.60	0.95	1.69
Final pH	5.18	5.60	5.95	6.69
ΔpH	**0.18**	**0.60**	**0.95**	**1.69**

Notice that ΔpH, the change in pH per milliequivalent of OH^- added, depends on the initial pH, with highest resistance to change at pH values close to the weak acid's pK_a. **Indeed, such weak acid-conjugate base combinations, called buffers, resist change most effectively when the desired pH falls within, ± 1.0 unit or less of their pK_a.**

Figure 2–6 also illustrates how the net charge on one molecule of a weak acid varies with pH. A fractional charge of −0.5 does not mean that an individual molecule bears a fractional charge but that the *probability* is 0.5 that a given molecule has a unit negative charge at any given moment in time. Consideration of the net charge on macromolecules as a function of pH provides the basis for separatory techniques such as ion exchange chromatography and electrophoresis (see Chapter 4).

The Propensity of a Proton to Dissociate Depends on Molecular Structure

Many acids of biologic interest possess more than one dissociating group. The presence of local negative charge hinders proton release from nearby acidic groups, raising their pK_a. This is illustrated by the pK_a values of the three dissociating groups of phosphoric acid and citric acid (**Table 2–2**). The effect of adjacent charge decreases with distance. The second pK_a for succinic acid, which has two methylene groups between its carboxyl groups, is 5.6, whereas the second pK_a for glutaric acid, which has one additional methylene group, is 5.4.

pK_a Values Depend on the Properties of the Medium

The pK_a of a functional group is also profoundly influenced by the surrounding medium. The medium may either raise or lower the pK_a relative to its value in water, depending on whether the undissociated acid or its conjugate base is the charged species. The effect of dielectric constant on pK_a may be observed by adding ethanol to water. The pK_a of a carboxylic acid *increases,* whereas that of an amine *decreases* on addition of ethanol because ethanol decreases the ability of water to solvate a charged species. The pK_a values of dissociating

TABLE 2–2 Relative Strengths of Monoprotic, Diprotic, and Triprotic Acids

Lactic acid	pK = 3.86		
Acetic acid	pK = 4.76		
Ammonium ion	pK = 9.25		
Carbonic acid	pK_1 = 6.37;	pK_2 = 10.25	
Succinic acid	pK_1 = 4.21;	pK_2 = 5.64	
Glutaric acid	pK_1 = 4.34;	pK_2 = 5.41	
Phosphoric acid	pK_1 = 2.15;	pK_2 = 6.82;	pK_3 = 12.38
Citric acid	pK_1 = 3.08;	pK_2 = 4.74;	pK_3 = 5.40

Note: Tabulated values are the pK_a values (–log of the dissociation constant).

groups in the interiors of proteins thus are profoundly affected by their local environment, including the presence or absence of water.

SUMMARY

- Water forms hydrogen-bonded clusters with itself and with other proton donors or acceptors. These extensive networks of hydrogen bonds account for the surface tension, viscosity, liquid state at room temperature, and solvent power of water.
- Compounds that contain O or N can serve as hydrogen bond donors and/or acceptors.
- Entropic forces dictate that amphipathic macromolecules bury nonpolar regions away from water.
- Salt bridges, hydrophobic interactions, and van der Waals forces participate in the formation of biomolecular complexes and maintenance of molecular conformation.
- pH is the negative log of $[H^+]$. A low pH characterizes an acidic solution, and a high pH denotes a basic solution.
- The strength of weak acids is expressed by pK_a, the negative log of the acid dissociation constant. Strong acids have low pK_a values and weak acids have high pK_a values.
- Buffers resist a change in pH when protons are produced or consumed. Maximum buffering capacity occurs within 1 pH unit on either side of pK_a. Physiologic buffers include bicarbonate, orthophosphate, and proteins.

REFERENCES

Reese KM: Whence came the symbol pH. Chem & Eng News 2004;82:64.

Segel IM: *Biochemical Calculations*. Wiley, 1968.

Skinner JL: Following the motions of water molecules in aqueous solutions. Science 2010;328:985.

Stillinger FH: Water revisited. Science 1980;209:451.

Suresh SJ, Naik VM: Hydrogen bond thermodynamic properties of water from dielectric constant data. J Chem Phys 2000;113:9727.

Wiggins PM: Role of water in some biological processes. Microbiol Rev 1990;54:432.

CHAPTER 3

Amino Acids & Peptides

Peter J. Kennelly, PhD, & Victor W. Rodwell, PhD

OBJECTIVES

After studying this chapter, you should be able to:

- Diagram the structures and write the three- and one-letter designations for each of the amino acids present in proteins.
- Provide examples of how each type of R group of the protein amino acids contributes to their chemical properties.
- List additional important functions of amino acids and explain how certain amino acids in plant seeds can severely impact human health.
- Name the ionizable groups of the protein amino acids and list their approximate pK_a values as free amino acids in aqueous solution.
- Calculate the pH of an unbuffered aqueous solution of a polyfunctional amino acid and the change in pH that occurs following the addition of a given quantity of strong acid or base.
- Define pI and explain its relationship to the net charge on a polyfunctional electrolyte.
- Explain how pK_a and pI can be used to predict the mobility at a given pH of a polyelectrolyte, such as an amino acid, in a direct-current electrical field.
- Describe the directionality, nomenclature, and primary structure of peptides.
- Describe the conformational consequences of the partial double-bond character of the peptide bond.

BIOMEDICAL IMPORTANCE

L-α-Amino acids provide the monomer units of the long polypeptide chains of proteins. In addition, these amino acids and their derivatives participate in such diverse cellular functions as nerve transmission and the biosynthesis of porphyrins, purines, pyrimidines, and urea. The neuroendocrine system employs short polymers of amino acids called *peptides* as hormones, hormone-releasing factors, neuromodulators, and neurotransmitters. Humans and other higher animals cannot synthesize 10 of the L-α-amino acids present in proteins in amounts adequate to support infant growth or to maintain adult health. Consequently, the human diet must contain adequate quantities of these *nutritionally essential* amino acids. Each day the kidneys filter over 50 g of free amino acids from the arterial renal blood. However, only traces of free amino acids normally appear in the urine because amino acids are almost totally reabsorbed in the proximal tubule, conserving them for protein synthesis and other vital functions.

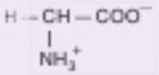
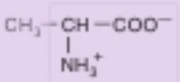
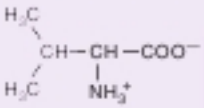
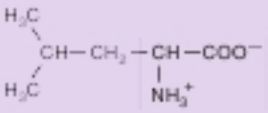
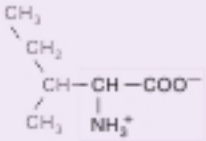
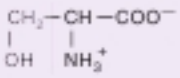
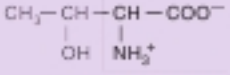
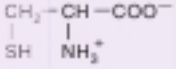
Certain microorganisms secrete free D-amino acids, or peptides that may contain both D- and L-α-amino acids. Several of these bacterial peptides are of therapeutic value, including the antibiotics bacitracin and gramicidin A, and the antitumor agent bleomycin. Certain other microbial peptides are, however, toxic. The cyanobacterial peptides microcystin and nodularin are lethal in large doses, while small quantities promote the formation of hepatic tumors. The ingestion of certain amino acids present in the seeds of legumes of the genus *Lathyrus* can result in lathyrism, a tragic irreversible disease in which individuals lose control of their limbs. Certain other plant seed amino acids have also been implicated in a neurodegenerative disease afflicting natives of Guam.

PROPERTIES OF AMINO ACIDS

The Genetic Code Specifies 20 L-α-Amino Acids

Although more than 300 amino acids occur in nature, proteins are synthesized almost exclusively from the set of 20 L-α-amino acids encoded by nucleotide triplets called **codons** (see Table 37–1). While the three-letter genetic code could potentially accommodate more than 20 amino acids, the genetic code is *redundant* since several amino acids are specified by multiple codons. Scientists frequently represent the sequences of peptides and proteins using one- and three-letter abbreviations for each amino acid **(Table 3–1)**. The R groups of amino acids may be either hydrophilic or hydrophobic **(Table 3–2)**; properties that affect their location in a protein's mature folded conformation (see Chapter 5). Some proteins contain additional amino acids that arise by the **posttranslational** modification of an amino acid already present in a peptide. Examples include the conversion of peptidyl proline and peptidyl lysine to 4-hydroxyproline and 5-hydroxylysine; the conversion of peptidyl glutamate to γ-carboxyglutamate; and the methylation, formylation, acetylation, prenylation, and phosphorylation of certain aminoacyl residues. These modifications significantly extend the functional diversity of proteins by altering their solubility, stability, catalytic activity, and interaction with other proteins.

TABLE 3–1 L-α-Amino Acids Present in Proteins

Name	Symbol	Structural Formula	pK_1	pK_2	pK_3
With Aliphatic Side Chains			**α-COOH**	**α-NH_2^+**	**R Group**
Glycine	Gly[G]		2.4	9.8	
Alanine	Ala[A]		2.4	9.9	
Valine	Val[V]		2.2	9.7	
Leucine	Leu[L]		2.3	9.7	
Isoleucine	Ile[T]		2.3	9.8	
With Side Chains Containing Hydroxylic(OH) Groups					
Serine	Ser[S]		2.2	9.2	about 13
Threonine	Thr[T]	 See below.	2.1	9.1	about 13
Tyrosine	Tyr[Y]	See below.			
With Side Chains Containing Sulfur Atoms					
Cysteine	Cys[C]		1.9	10.8	8.3

(*Continued*)

TABLE 3–1 L-α-Amino Acids Present in Proteins (Continued)

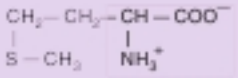
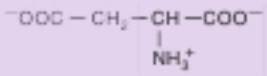
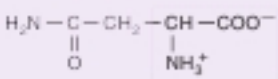
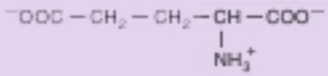
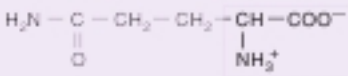
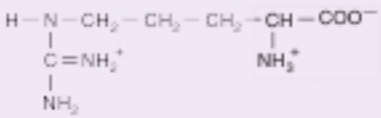
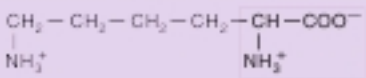
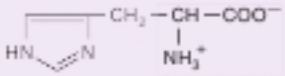
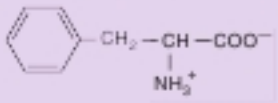
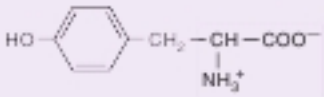
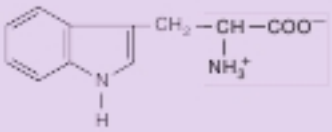
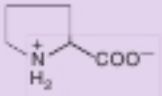
Name	Symbol	Structural Formula	pK_1	pK_2	pK_3
With Side Chains Containing Sulfur Atoms					
Methionine	Met[M]		2.1	9.3	
With Side Chains Containing Acidic Groups or Their Amides					
Aspartic Acid	Asp[D]		2.1	9.9	3.9
Asparagine	Asn[N]		2.1	8.8	
Glutamic Acide	Glu[E]		2.1	9.5	4.1
Glutamine	Gin[Q]		2.2	9.1	
With Side Chains Containing Basic Groups					
Argine	Arg[R]		1.8	9.0	12.5
Lysine	Lys[K]		2.2	9.2	10.8
Histidine	His[H]		1.8	9.3	6.0
Containing Aromatic Rings					
Histidine	His[H]	See above			
Phenylalanine	Phe[F]	See above. 	2.2	9.2	
Tyrosine	Try[Y]		2.2	9.1	10.1
Tryptophan	Trp[W]		2.4	9.4	
Imino Acid					
Proline	Pro[P]		2.0	10.6	

TABLE 3–2 Hydrophilic & Hydrophobic Amino Acids

Hydrophilic	Hydrophobic
Arginine	Alanine
Asparagine	Isoleucine
Aspartic acid	Leucine
Cysteine	Methionine
Glutamic acid	Phenylalanine
Glutamine	Proline
Glycine	Tryptophan
Histidine	Tyrosine
Lysine	Valine
Serine	
Threonine	

The distinction is based on the tendency of their R groups to associate with, or to minimize contact with, an aqueous environment.

Selenocysteine, the 21st Protein L-α-Amino Acid

Selenocysteine **(Figure 3–1)** is an L-α-amino acid present in proteins from every domain of life. Humans contain approximately two dozen selenoproteins that include certain peroxidases and reductases, selenoprotein P, which circulates in the plasma, and the iodothyronine deiodinases responsible for converting the prohormone thyroxine (T_4) to the thyroid hormone 3,3′, 5-triiodothyronine (T_3) (see Chapter 41). Since peptidyl selenocysteine is inserted directly into a growing polypeptide during *translation,* it is commonly termed the "21st amino acid." However, unlike the other 20 protein amino acids, incorporation of selenocysteine is specified by a large and complex genetic element for the unusual tRNA called $tRNA^{Sec}$ which utilizes the UGA anticodon that normally signals STOP rather than a simple triplet codon. Rather, the protein synthetic apparatus recognizes a UGA codon as coding for selenocysteine when it is accompanied by a stem-loop structure, the selenocysteine insertion element, in the untranslated region of the mRNA (see Chapter 27).

Stereochemistry of the Protein Amino Acids

With the sole exception of glycine, the α-carbon of every amino acid is chiral. Although some protein amino acids are dextrorotatory and some levorotatory, all share the absolute

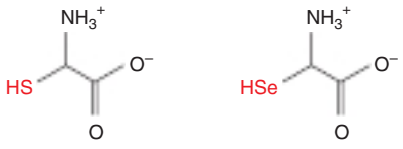


FIGURE 3–1 Cysteine (left) and selenocysteine (right). pK_3, for the selenyl proton of selenocysteine is 5.2. Since this is 3 pH units lower than that of cysteine, selenocysteine represents a better nucleophile at or below pH 7.4.

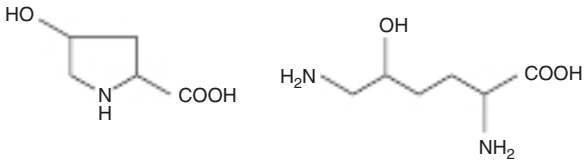


FIGURE 3–2 4-Hydroxyproline and 5-hydroxylysine.

configuration of L-glyceraldehyde and thus are defined as L-α-amino acids. Even though almost all protein amino acids are (*S*), the failure to use (*R*) or (*S*) to express *absolute* stereochemistry is no mere historical aberration. L-Cysteine is (*R*) since the atomic mass of the sulfur atom on C3 exceeds that of the amino group on C2. More significantly, in mammals the biochemical reactions of L-α-amino acids, their precursors, and their catabolites are catalyzed by enzymes that act exclusively on L-isomers, irrespective of their absolute configuration.

Posttranslational Modifications Confer Additional Properties

While some prokaryotes incorporate pyrrolysine into proteins, and plants can incorporate azetidine-2-carboxylic acid, an analog of proline, a set of just 21 L-α-amino acids clearly suffices for the formation of most proteins. Posttranslational modifications can, however, generate novel R groups that impart further properties. In collagen, protein-bound proline and lysine residues are converted to 4-hydroxyproline and 5-hydroxylysine **(Figure 3–2)**. The carboxylation of glutamyl residues of proteins of the blood coagulation cascade to γ-carboxyglutamyl residues **(Figure 3–3)** completes a site for chelating the calcium ion essential for blood coagulation. The amino acid side chains of histones are subject to numerous modifications, including acetylation and methylation of lysine and methylation and deimination of arginine (see Chapters 35 and 38). It is also now possible in the laboratory to genetically introduce many different unnatural amino acids into proteins, generating proteins via recombinant gene expression with new or enhanced properties and providing a new way to explore protein structure–function relationships.

Extraterrestrial Amino Acids Have Been Detected in Meteorites

The existence of extraterrestrial amino acids was first reported in 1969 following analysis of the famous Murchison meteorite from southeastern Australia. The presence of amino acids in other meteorites, including some pristine examples from Antarctica, has now been amply confirmed. Unlike the amino acids synthesized by terrestrial organisms, these meteorites

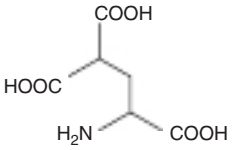


FIGURE 3–3 γ-Carboxyglutamic acid.

contain racemic mixtures of D- and L-isomers of multiple protein amino acids as well as biologically important nonprotein α-amino acids such as *N*-methylglycine (sarcosine) and β-alanine. Several novel amino acids that lack terrestrial counterparts of biotic origin were also discovered. Nucleobases, activated phosphates, and molecules related to sugars have also been detected in meteorites. These findings offer potential insights into the prebiotic chemistry of earth, and impact the search for extraterrestrial life. Some speculate that meteorites may have contributed to the origin of life on our planet, a conjecture entitled Panspermia, by delivering extraterrestrially generated organic molecules or even intact microorganisms to our earth.

L-α-Amino Acids Serve Additional Metabolic Roles

L-α-Amino acids fulfill vital metabolic roles in addition to serving as the "building blocks" of proteins. For example, ornithine and citrulline are key intermediates in the urea cycle (see Figure 28–16), while S-adenosyl-methionine serves as a methyl-group donor for many enzyme-catalyzed reactions. Tyrosine is a precursor of thyroid hormone, while both tyrosine and phenylalanine are metabolized to produce epinephrine, norepinephrine, and dihydroxyphenylalanine (DOPA). Glutamate is both a neurotransmitter as well as a precursor of a second neurotransmitter, γ-aminobutyric acid (GABA).

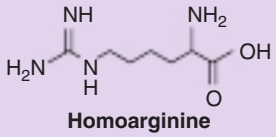
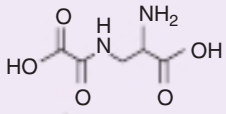
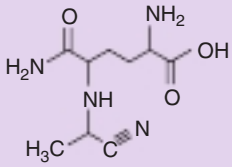
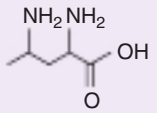
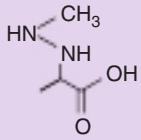
Certain Plant L-α-Amino Acids Can Adversely Impact Human Health

The consumption of plants that contain certain nonprotein amino acids can adversely impact human health. The seeds and seed products of three species of the legume *Lathyrus* have been implicated in the genesis of **neurolathyrism**, a profound neurologic disorder characterized by progressive and irreversible spastic paralysis of the legs. Lathyrism occurs widely during famines, when *Lathyrus* seeds may become a major part of the diet. L-α-Amino acids that have been implicated in human neurologic disorders, notably neurolathyrisms, include L-homoarginine and β-*N*-oxalyl-L-α,β-diaminopropionic acid (β-ODAP **Table 3–3**). The seeds of another *Lathyrus* legume, the "sweet pea," contain the osteolathyrogen γ-glutamyl-β-aminopropionitrile (BAPN), a glutamine derivative of β-aminopropionitrile (structure not shown). The seeds of certain *Lathyrus* species also contain α,γ-diaminobutyric acid, an analog of ornithine that inhibits the hepatic urea cycle enzyme ornithine transcarbamylase, leading to ammonia toxicity. Finally, L-β-methylaminoalanine, a neurotoxic amino acid that is present in *Cycad* seeds, has been implicated as a risk factor for neurodegenerative diseases including amyotrophic lateral sclerosis–Parkinson dementia complex in natives of Guam who consume either fruit bats that feed on cycad fruit, or flour made from cycad seeds.

D-Amino Acids

D-Amino acids occur naturally throughout the biosphere, including free D-serine and D-aspartate in human brain tissue,

TABLE 3–3 Potentially Toxic L-α-Amino Acids

Nonprotein L-α-Amino Acid	Medical Relevance
 Homoarginine	Cleaved by arginase to L-lysine and urea. Implicated in human neurolathyrism.
 β-*N*-Oxalyl diaminopropionic acid (β-ODAP)	A neurotoxin. Implicated in human neurolathyrism.
 β-*N*-Glutamylamino-propiononitrile (BAPN)	An osteolathyrogen.
 2,4-Diaminobutyric acid	Inhibits ornithine transcarbamylase, resulting in ammonia toxicity.
 β-Methylaminoalanine	Possible risk factor for neurodegenerative diseases.

D-alanine and D-glutamate in the cell walls of gram-positive bacteria, and D-amino acids in certain peptides and antibiotics produced by bacteria, fungi, reptiles, and amphibians. *Bacillus subtilis* excretes D-methionine, D-tyrosine, D-leucine, and D-tryptophan to trigger biofilm disassembly, and *Vibrio cholerae* incorporates D-leucine and D-methionine into the peptide component of its peptidoglycan layer.

PROPERTIES OF THE FUNCTIONAL GROUPS OF AMINO ACIDS

Amino Acids May Have Positive, Negative, or Zero Net Charge

In aqueous solution, the charged and uncharged forms of the ionizable weak acid groups —COOH and $—NH_3^+$ exist in dynamic protonic equilibrium:

$$R—COOH \rightleftarrows R—COO^- + H^+$$
$$R—NH_3^+ \rightleftarrows R—NH_2 + H^+$$

While both R—COOH and $R—NH_3^+$ are weak acids, R—COOH is a far stronger acid than $R—NH_3^+$. Thus, at physiologic pH