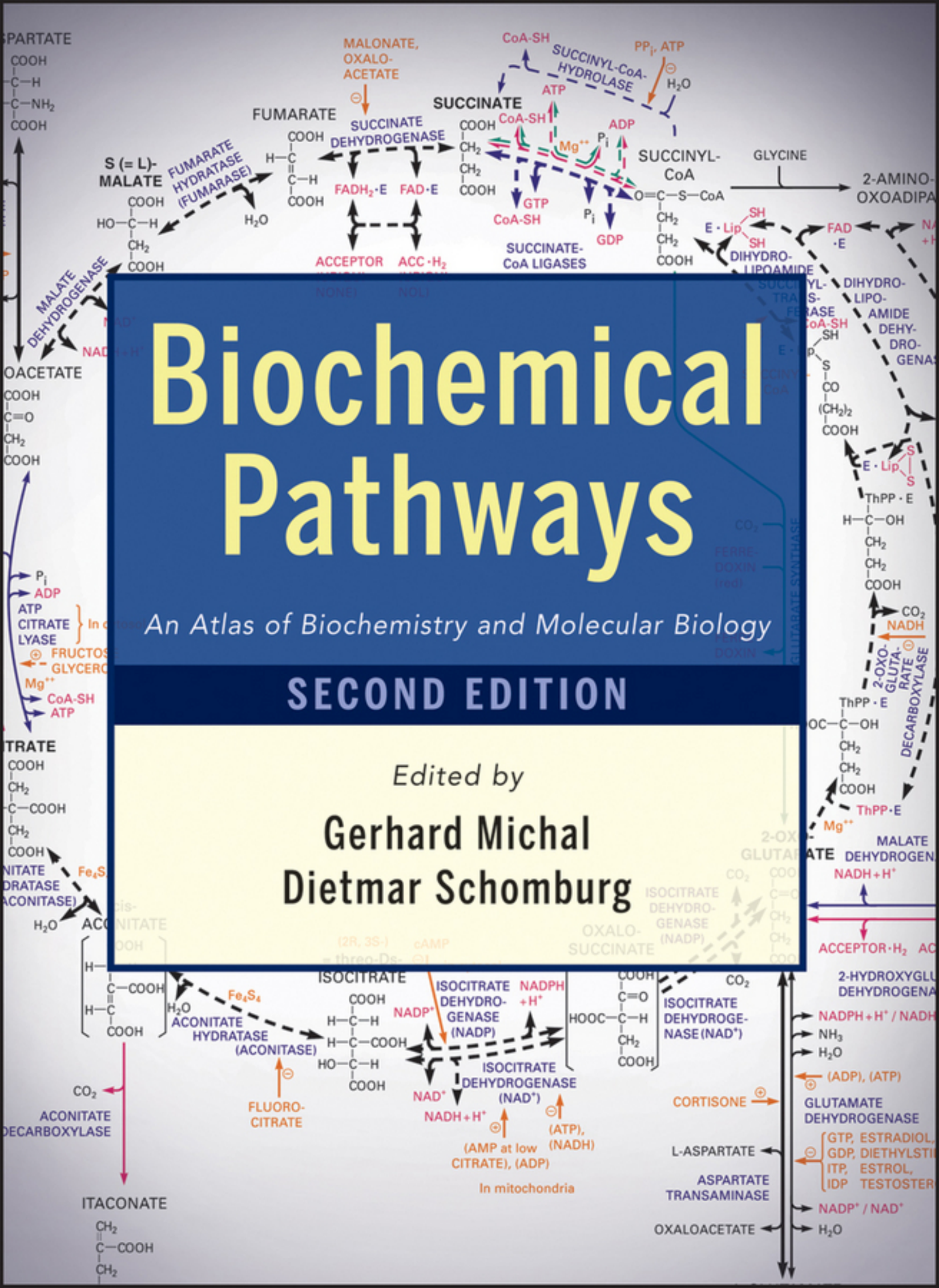

Biochemical Pathways
An Atlas of Biochemistry and Molecular Biology
SECOND EDITION
Edited by
Gerhard Michal
Dietmar Schomburg

BIOCHEMICAL PATHWAYS:
AN ATLAS OF BIOCHEMISTRY AND MOLECULAR BIOLOGY

Biochemical Pathways Posters available from Roche Applied Science

Gerhard Michal's famous biochemical pathways posters have been a valuable resource for the global biochemistry community since 1968. Updated and revised, the *Biochemical Pathways Wallcharts* are an ideal companion to this text. Paired together, the book and wallcharts are perfect for researchers and students in biochemistry, biology, medicine, and physiology.

The wallcharts are also a great gift for anyone interested in following the myriad chemical reactions in our cells.To obtain this pair of large, detailed wall charts, contact Roche at https://www.roche-applied-science.com/techresources/index.jsp

BIOCHEMICAL PATHWAYS:

An Atlas of Biochemistry and Molecular Biology

Second Edition

Edited by

Gerhard Michal

Dietmar Schomburg

A JOHN WILEY & SONS, INC., PUBLICATION

Published by John Wiley & Sons, Inc., Hoboken, New Jersey
Published simultaneously in Canada

Library of Congress Cataloging-in-Publication Data:

Biochemical pathways : an atlas of biochemistry and molecular biology / edited by Dietmar Schomburg, Gerhard Michal. -- 2nd ed.
p. cm.
Includes bibliographical references and index.
ISBN 978-0-470-14684-2
1. Metabolism--Atlases. I. Schomburg, D. (Dietmar) II. Michal, Gerhard.
QP171.B685 2012
612.3'9--dc23

2011041441

Printed in the Singapore
10 9 8 7 6 5 4 3 2 1

Contents

Preface to the Second Edition

Since the publication of the first edition of 'Biochemical Pathways' in 1999 the molecular life sciences (encompassing biology, biochemistry, pharmacy and medicine) have undergone dramatic changes. With the extremely rapid development in the 'OMICS' analytical techniques (Genomics, transcriptomics, proteomics, metabolimics) we are in principle able to determine the genome of a microorganism in one day and a human genome for a couple of thousand dollars. We have also seen the advent of 'systems biology', which, based on the measured OMICS-data, aims at analysis and even prediction of biological functions by the construction of computer models. These models simulate the reaction of biological systems, including whole cells, to changes in the environment, genetic disorders or mutations.

Based on the annotation of the genome and experimental data, metabolic, regulatory and signal transduction pathways and networks are constructed and mathematically formulated. They depend entirely on our knowledge of biochemical pathways, as they are presented in this book.

As outlined in the preface to the first edition, one of us (GM) began early in the 1960s to combine an extract of the biochemical knowledge in a wall chart. The other of us (DS), towards the end of his student life saw the 'Biochemical Pathways' wall chart or 'Boehringer chart' in almost every lab working in the field of biochemistry or molecular biology. (At present, it is distributed as the 4th edition by Roche Diagnostics GmbH, Mannheim). He was impressed by the puzzle work biochemists had performed for almost one century. This presentation of important features of biochemistry was extended in the first edition of the 'Biochemical Pathways' book, which has become the standard book of reference in his and many other labs since then. In its focus on pathways and networks it is unique and was published exactly at a time when pathways, networks and systems became the focus of biochemical research. These areas have become the major fields of DS's research work in the last decade.

The fields of activities on both sides encouraged us to combine our experiences in writing and publishing the second edition of this book. The task became larger than expected on the first glance. Since the publication of the first edition our knowledge has increased tremendously. The selection of the facts to be dealt with and their condensation into a short, but legible form was no easy task. We could persuade expert authors to help us with the book. We both had a highly enjoyable cooperation and could now finally finish this work. We want to thank all authors for their contributions. In addition, Robbe Wünschiers likes to express his gratitude to Dr. Rainer Lemke for supporting the revision of the chapters.

The book not only gained one half in volume, but every sentence and every figure had to be checked and often modified. More than half of the many hundreds of figures in the book had to replaced, modified or added in this second edition.

We hope that it will help students and researchers to obtain a deeper understanding of the pathways and networks that determine biological functions.

Gerhard Michal
Dietmar Schomburg

From the Preface to the First Edition

This book is not intended to be a textbook of biochemistry in the conventional sense. There is no shortage of good biochemistry textbooks. which outline how biochemical knowledge has been gained, trace the logical and experimental developments in this field and present advances in their historical sequence.

In contrast, this book tries to condense important aspects of current knowledge. Its goal is to give concise information on the metabolic sequences in the pathways, the chemistry and enzymology of conversions, the regulation of turnover and the effect of disorders. This concentration on the sequence of facts has entailed the omission of researchers' names, experimental methods and the discussion of how results have been obtained. For information on these aspects, and for an introduction to the fundamentals of biological science, it is necessary to consult textbooks.

The scope of this book is general biochemistry, encompassing bacteria (and to some extent archaea), plants, yeasts and animals. Although a balanced representation is intended, personal interest naturally plays a role in the selection of topics. In a number of cases, the chemistry of the reactions is given in more detail, especially at metabolic key and branching points. Human metabolism, its regulation and disorders as a result of disease is a frequent topic. On the other hand, some chapters are especially devoted to bacterial metabolism.

This book grew out of my interest in metabolic interrelationships and regulation which was stimulated by my professional work at Boehringer Mannheim GmbH, Germany. Previously, this interest led me to compile the 'Biochemical Pathways'wall chart, the first edition of which appeared 40 years ago. Three more editions followed, which have been widely distributed. As a result of this experience, I developed a preference for the graphic presentation of scientific facts. In contrast to texts, illustrations allow the simultaneous display of different aspects, such as structural formulas, enzyme catalysis and its regulation, the involvement of cofactors, the occurrence of enzymes in various kingdoms of biology, etc. This form of presentation facilitates a rapid overview. A standard set of conventions is used in all illustrations (representation of formulas, symbols for proteins, the use of colors, the shape of arrows, etc. - the rare exceptions are indicated), and this assists in finding the facts quickly.

Tables have been added to provide more detailed information. They list additional properties of the system components, homologies, etc. The text plays only a supportive role. It gives a concise description of the reactions and their regulation, and puts them into the general metabolic context.

In many cases, current knowledge focuses on a limited in number of species. A rough classification of the occurrence of pathways is given by the color or the reaction arrows in the illustrations, but both generalizations and specialization are expected to be found in the future, which will necessitate modification of the picture.

The literature references have been limited in number and they usually cite recent review articles and books, if possible, from readily accessible sources. They were selected to provide more detailed information on new developments and additional references for the interested reader. There are no references to long-established biochemical facts which can be found in any textbook. I hope that this restriction will be acceptable to readers, since a complete listing of all sources for the statements presented here would take up a major portion of this volume. To compensate for the omission of such general references, a special chapter on electronic data banks and major printed sources has been added at the end of the book.

Most of all I want to thank my wife Dea, who has often encouraged me during the long time required to fiish this work. She has given me valuable advice and support in checking the text of the English edition. Without her understanding and her help this book would not have been brought to completion.

Gerhard Michal

Contributors

Helmut Burtscher, Roche Diagnostics GmbH, D-82372 Penzberg

Antje Chang, Enzymeta GmbH, D-50374 Erftstadt

Petra Dersch, Helmholtz Center for Infection Research, Dept. of Molecular Infection Biology, D-38124 Braunschweig

Julia Garbe, Institute for Microbiology, Technische Universität, D-38106 Braunschweig

Anton Haselbeck, MAB Discovery GmbH, D-82061 Neuried

Elmar Heinzle, Technical Biochemistry, Universität des Saarlandes, D-66123 Saarbrücken

Dieter Jahn, Institute for Microbiology, Technische Universität, D-38106 Braunschweig

Martina Jahn, Institute for Microbiology, Technische Universität, D-38106 Braunschweig

Wilhelm Just, Biochemistry Center, University, D-69120 Heidelberg

Horst Klima, Roche Diagnostics GmbH, D-82372 Penzberg

Klaus Klumpp, Hoffmann-La Roche Inc., Nutley NJ 07110

Gerhard Michal, Roche Diagnostics GmbH, D-82372 Penzberg, formerly Boehringer Mannheim GmbH (ret.)

Peter Müller, Helmholtz Centre for Infection Research, Dept. Gene Regulation & Differentiation, D-38124 Braunschweig

Gerhard Niederfellner, Roche Diagnostics GmbH, D-82372 Penzberg

Dieter Oesterheldt, Max-Planck Institute for Biochemistry, D-82152 Martinsried

Susanne Peifer, Technical Biochemistry, Universität des Saarlandes, D-66123 Saarbrücken

Ernst Peter Rieber, Institute for Immunology, Technische Universität, D-01011 Dresden

Stefan Ries, Roche Diagnostics GmbH, D-82372 Penzberg

Max Schobert, Institute for Microbiology, Technische Universität, D-38106 Braunschweig

Dietmar Schomburg, Institute of Biochemistry, Biotechnology & Bioinformatics, Technische Universität, D-38106 Braunschweig

Ida Schomburg, Enzymeta GmbH, D-50374 Erftstadt

Annika Steen, Institute for Microbiology, Technische Universität, D-38106 Braunschweig

Josef Wachtveitl, Institute for Physical and Theoretical Chemistry, University, D-60438 Frankfurt/M.

Röbbe Wünschiers, Biotechnology/Computational Biology, University of Applied Sciences, D-09648 Mittweida

1 Introduction and General Aspects

Gerhard Michal and Dietmar Schomburg

1.1 Organization of This Book

This book deals with the chemistry of living organisms. However, this topic cannot be considered in an isolated way, but has to be placed into a more general context. In two introductory chapters, a short outline of interconnections with neighboring sciences is given.

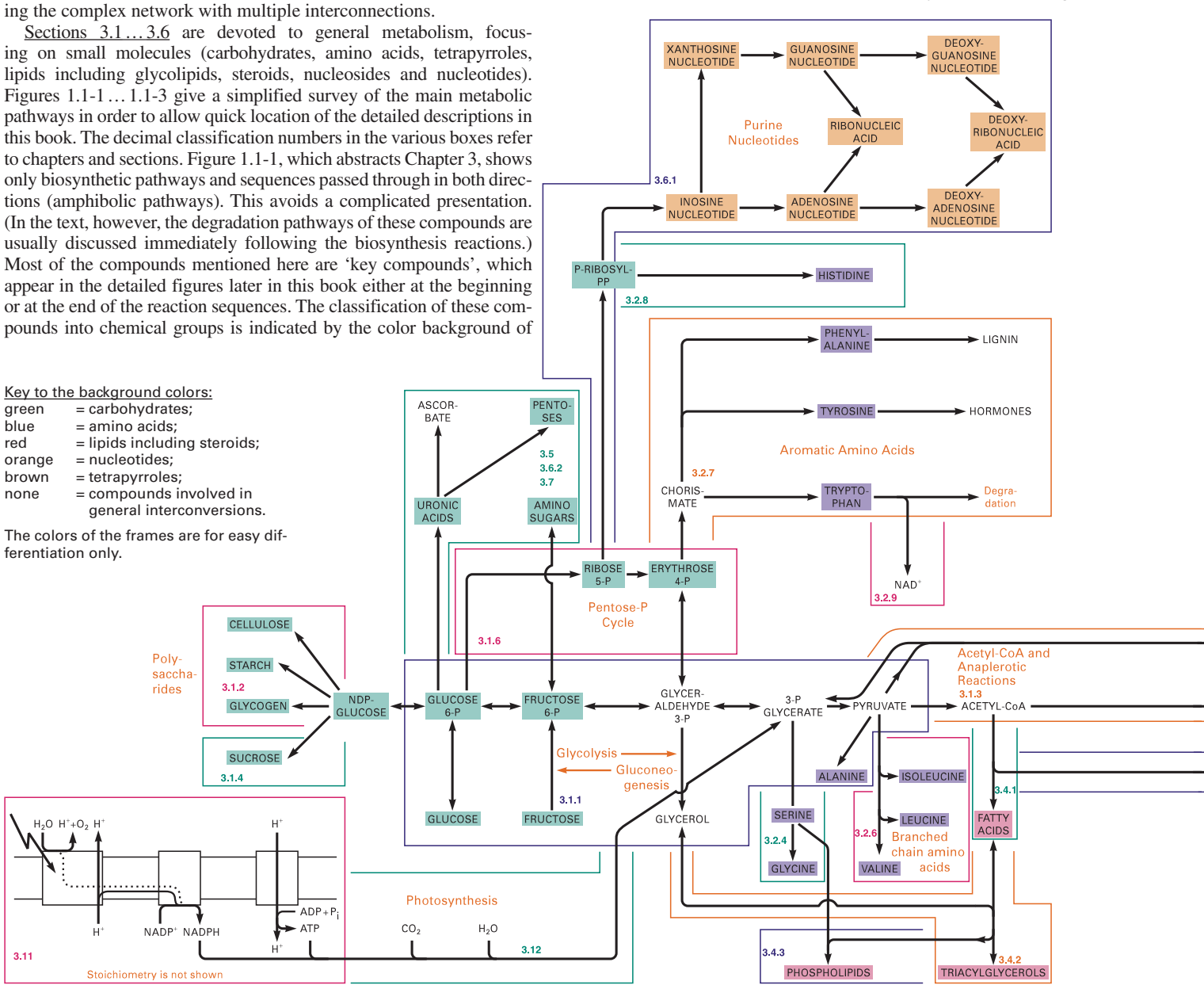
Chapter 1 deals with the organic chemistry of important components present in living organisms and with the physical chemistry of reactions.

Chapter 2 describes the overall organization of cells and their organelles as well as the structure of proteins and nucleic acids. This is followed by a discussion of enzyme function, which depends on the protein structure and regulates almost all biological processes.

The topics of **Chapter 3** are various aspects of metabolism, showing the complex network with multiple interconnections.

Sections 3.1 … 3.6 are devoted to general metabolism, focusing on small molecules (carbohydrates, amino acids, tetrapyrroles, lipids including glycolipids, steroids, nucleosides and nucleotides). Figures 1.1-1 … 1.1-3 give a simplified survey of the main metabolic pathways in order to allow quick location of the detailed descriptions in this book. The decimal classification numbers in the various boxes refer to chapters and sections. Figure 1.1-1, which abstracts Chapter 3, shows only biosynthetic pathways and sequences passed through in both directions (amphibolic pathways). This avoids a complicated presentation. (In the text, however, the degradation pathways of these compounds are usually discussed immediately following the biosynthesis reactions.) Most of the compounds mentioned here are 'key compounds', which appear in the detailed figures later in this book either at the beginning or at the end of the reaction sequences. The classification of these compounds into chemical groups is indicated by the color background of the names. Section 3.7 deals with cofactors and vitamins, which are involved in many reactions of general metabolism. Sections 3.8 and 3.9 describe the metabolism of DNA in bacteria and eukarya and the repair systems of these essential information carriers. The special metabolism of bacteria (including energy aspects), the biosynthesis and the effects of antibiotics are topics of Section 3.10. Aerobic respiration and its central role in energy turnover, as well as the photosynthetic reactions that are the source of almost all compounds in living beings, are discussed in Sections 3.11 and 3.12. Many special metabolic reactions take place in plants. These are summarized in Section 3.13.

The biosynthesis of proteins in bacteria and eukarya, and their consecutive modification, as well as the cell cycle, are discussed in **Chapter 4**. Figure 1.1-2 gives a short outline of these reactions, subdivided into bacterial reactions (left) and eukaryotic reactions (right).



Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology, Second Edition. Edited by Gerhard Michal and Dietmar Schomburg.

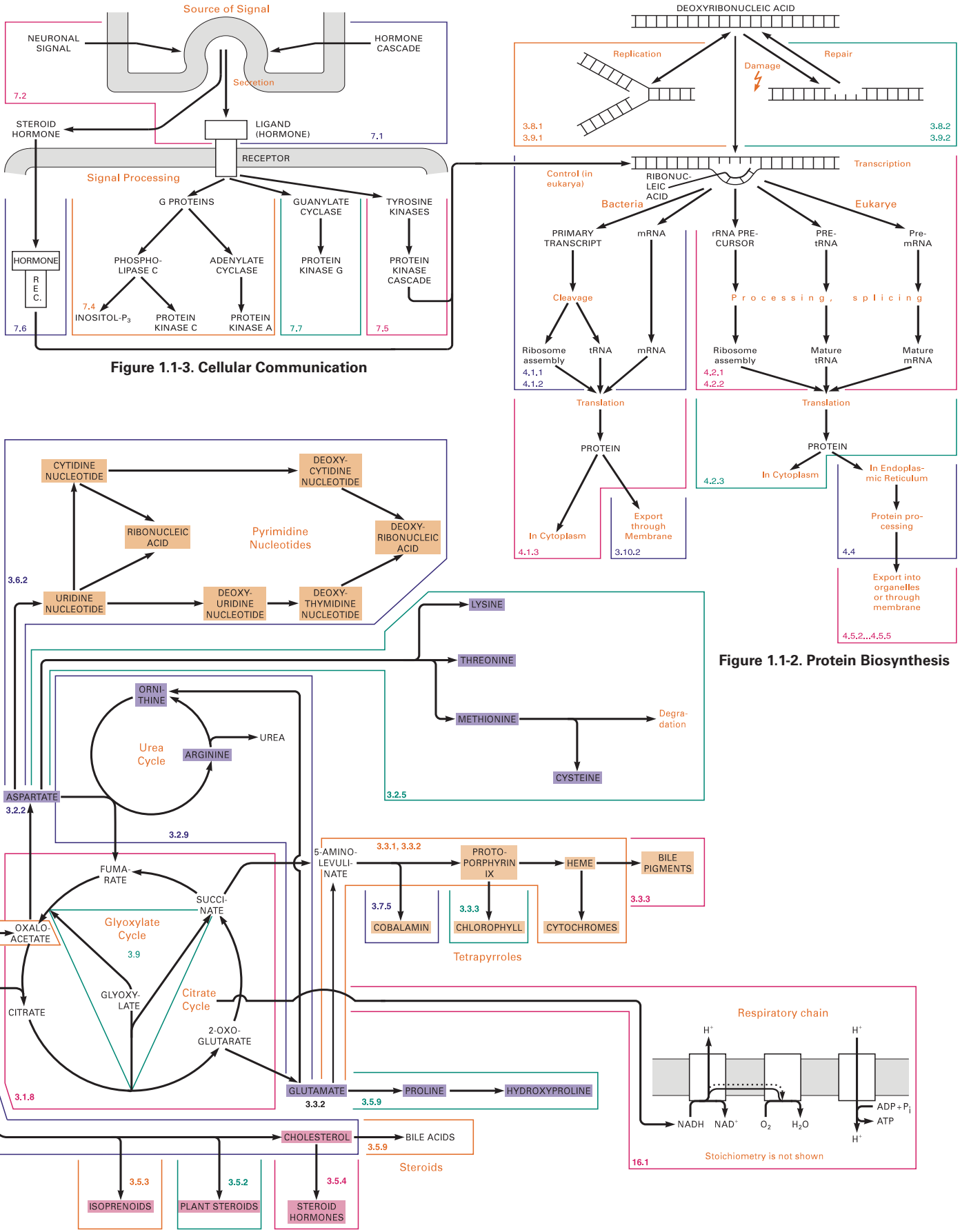


Figure 1.1-3. Cellular Communication

Figure 1.1-2. Protein Biosynthesis

Figure 1.1-1. Biosynthetic Reactions in General Metabolism

Viruses, which utilize these mechanisms in hosts, are discussed in **Chapter 5**.

Chapter 6 gives a survey of transport mechanisms through membranes and within vessels.

The topic of **Chapter 7** is cellular communication and the regulation mechanisms employed by multicellular organisms. Figure 1.1-3 briefly summarizes these multiple interconnections.

Chapter 8 deals with the defense mechanisms of higher animals and **Chapter 9** with blood coagulation.

Every presentation can only contain a selection of the present knowledge. For this reason, the final **Chapter 10,** is intended to assist in obtaining further information from electronic sources, which offer the most comprehensive collection of scientific results available today.

1.1.1 Conventions Used in This Book

1. A decimal classification system is used throughout with the following subdivisions: chapters, sections, subsections. Figures, tables, and formulas are assigned to the relevant sections, e.g., Figure 3.7.6-1.

Reactions:

2. Whenever available, the "Accepted Names" as defined by the IUMB Biochemical Nomenclature Committee are used for enzymes and substrates. The enzyme classification scheme (EC numbers) and the transporter classification scheme (TC numbers) are listed in the index.
3. Substrates of enzymatic reactions are printed in black, enzymes in blue, coenzymes in red. Regulatory effects are shown in orange. This color is also used for pathway names and for information on the location of a reaction. For numbering systems, green is used.
4. The color of the reaction arrows shows where the reaction was observed (or at least where reasonable indications for its occurrence exist): black = general pathway, blue = in animals, green = in plants and yeasts, red = in prokarya (bacteria and archaea).
5. Bold arrows indicate main pathways of metabolism.
6. Points on both ends of an arrow ↔ indicate noticeable reversibility of this reaction under biological conditions. Unless expressly noted, this type of arrow does not indicate mesomeric (resonance stabilized) states of a compound, contrary to usage in organic chemistry.
7. Double arrows ⇌ are used when the interconversion of two compounds proceeds via different reactions in each direction (e.g., for some steps of glycolysis).
8. Dashed reaction arrows show conversions with primarily catabolic (degradative) importance. Full line arrows show either mainly anabolic (biosynthetic) reactions or reactions in biological systems which are frequently passed through in both directions (amphibolic reactions).

Regulation:

9. Necessary cofactors, activating ions etc. are printed in orange next to a reaction arrow.
10. Full line orange arrows with an accompanying ⊕ or ⊖ indicate that the respective factor exerts 'fast' activation or inhibition of the reaction (by allosteric mechanisms, product inhibition etc.). Dashed arrows are used if the amount of enzyme protein is regulated, e.g., by varied expression or by changes in the degradation rate. If only one of multiple enzymes is regulated in this way, it is indicated by Roman numbers.

Enzymes and Proteins:

11. When enzyme complexes are involved, the respective components are schematically drawn in blue-lined boxes with rounded edges. This does not express the spatial structure. If possible, interacting components are drawn next to each other.
12. When a sequence of domains occurs in a protein, special symbols are used for the individual domains. They are explained next to the drawing.
13. When the peptide chain has to be shown, helices are drawn as  (e.g., in transmembrane domains), otherwise they are symbolized as .

Abbreviations and Notations:

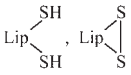
14. Organic phosphate is generally abbreviated as –P, inorganic phosphate and pyrophosphate as P_i and PP_i respectively. In drawings where the reaction mechanism is emphasized, phosphate residues are shown as $–O–PO_3^{2–}$.
15. Braces {} are used for atoms or residues which formally enter or leave during a reaction, if the molecular context is unknown.
16. While notations for genes are usually printed with small case letters (e.g., raf), the respective gene products (proteins) are written with a capitalized first letter (e.g., Raf). A number of proteins are defined by their molecular mass in kDa, e.g., p53.
17. When protein names are abbreviated, the notation frequently uses capitalized letters, e.g., cyclin dependent kinases = CDK in accordance with the literature.
18. A list of common abbreviations used throughout the book is given in 1.1.2. Less frequently used abbreviations are defined in the text.

Literature:

19. Only some recent references, primarily review articles and monographs, are listed at the end of the various sections. For more details refer to the literature quoted in these references, to electronic data banks, to review books and journals and to biochemistry textbooks.
20. Chapter 10 contains a survey on electronic data banks and a list of printed sources, which have been used frequently during the writing of this book.

1.1.2 Common Abbreviations (Other abbreviations are defined in the text)

aa	Amino acid
Acc, $AccH_2$	Acceptor, reduced acceptor (unspecified)
ACP	Acyl carrier protein
ATP, ADP, AMP, A	Adenosine tri-, di-, monophosphate, adenosine
bp, kbp	base pair (in DNA), kilobase pairs
cAMP	Cyclic AMP = adenosine 3,5-monophosphate
cGMP	Cyclic GMP = guanosine 3,5-monophosphate
CoA-SH, CoA-S-	Coenzyme A
CTP, CDP, CMP, C	Cytidine tri-, di-, monophosphate, cytidine
Cyt	Cytochrome
Da, kDa	Dalton, kilodalton (unit of molecular mass)
dATP, dADP, dAMP, dA	Deoxyadenosine tri-, di-, monophosphate, deoxyadenosine
dCTP, dCDP, dCMP, dC	Deoxycytidine tri-, di-, monophosphate, deoxycytidine
dGTP, dGDP, dGMP, dG	Deoxyguanosine tri-, di-, monophosphate, deoxyguanosine
dTTP, dTDP, dTMP, dT	Deoxythymidine tri-, di-, monophosphate, deoxythymidine
DNA	Deoxyribonucleic acid
E	Enzyme
EC number	Enzyme classification according to the IUBMB EC classification
ER	Endoplasmatic reticulum
ETF	Electron transferring flavoprotein
F_{430}	A corrinoid coenzyme (Ni)
FAD, $FADH_2$	Flavin-adenine dinucleotide, reduced flavin-adenine dinucleotide
Fd	Ferredoxin
FMN, $FMNH_2$	Flavin mononucleotide, reduced flavin mononucleotide
Fp	Flavoprotein
ΔG	Change of free energy (see 1.5.1)
G6P	Glucose 6-phosphate
GSH, GSSG	Glutathione, oxidized glutathione
GTP, GDP, GMP, G	Guanosine tri-, di-, monophosphate, guanosine
Ig	Immunoglobulin
ITP, IDP, IMP, I	Inosine tri-, di-, monophosphate, inosine
k	Velocity constant of a reaction (1.5.4)
K	Equilibrium constant of a reaction (see 1.5.1)
K_S, K_I, K_D	Dissociation constants (see 1.5.4, 7.1.2)
K_M	Michaelis constant (see 1.5.4)

kb	Kilobases (10^3 bases)
λ	Wavelength of light
	α-Lipoic acid, oxidized α-lipoic acid
NAD^+, NADH + H^+	Nicotinamide-adenine dinucleotide, reduced nicotinamide-adenine dinucleotide
$NADP^+$, NADPH + H^+	Nicotinamide-adenine dinucleotide phosphate, reduced nicotinamide-adenine dinucleotide phosphate
nt	Nucleotide
NTP, NDP, NMP, N	Any nucleotide tri-, di-, monophosphate or nucleoside
PAP	Adenosine 3′,5′-diphosphate
PAPS	3′-Phosphoadenylylsulfate
PEP	Phosphoenolpyruvate
P_i, Pp_i	Inorganic phosphate, inorganic pyrophosphate
pH	Negative decadic logarithm of the H^+ concentration
pK	Negative decadic logarithm of a dissociation constant
PQQ	Pyrroloquinoline quinone
PRPP	α-D-5-Phosphoribosylpyrophosphate
PyrP	Pyridoxal phosphate
RNA	Ribonucleic acid
mRNA, rRNA, tRNA	Messenger-, ribosomal-, transfer ribonucleic acid
R-S-S-R	Disulfide group of amino acids or peptides
S	Svedberg units (sedimentation coefficient)
SAH	S-Adenosylhomocysteine
SAM	S-Adenosylmethionine
THF	5,6,7,8-Tetrahydrofolate
THMPT	5,6,7,8-Tetrahydromethanopterin
ThPP	Thiamin pyrophosphate
UDPG	Uridine diphosphate glucose
UQ, UQH_2	Ubiquinone, reduced ubiquinone
UTP, UDP, UMP, U	Uridine tri-, di-, monophosphate, uridine

Abbreviations for amino acids are listed in Figure 1.3.2, abbreviations for sugars in Figure 4.4.1-1.

1.2 Carbohydrate Chemistry and Structure

Carbohydrate monomers are of the general formula $(CH_2O)_n$. They have the chemical structure of aldehydes or ketones with multiple hydroxyl groups (aldoses and ketoses, respectively). A common name of monomers and dimers is 'sugar'.

The large number of reactive groups, together with the stereoisomers causes a multiplicity of structures and reaction possibilities. Besides 'pure' carbohydrate monomers, oligomers (3.1.4) and polymers (3.1.2), carboxylic (3.1.5.1…2) and amino (3.1.7) derivatives, polyalcohols (3.1.5.5), deoxy sugars (3.1.5.6) etc., exist in nature.

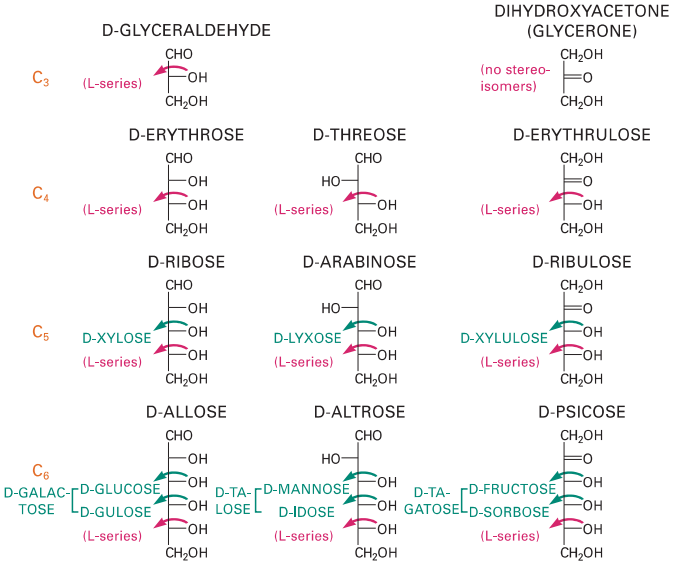


Figure 1.2-1. Nomenclature of Carbohydrates
The compounds printed in green are formally obtained by epimerization at the indicated positions. The L-enantiomers are the mirror images at the perpendicular mirror plane.

Carbohydrates are the primary products of photosynthesis (3.12.2) and function as energy storage forms (e.g., starch, glycogen, 3.1.2), as part of nucleic acid and nucleotide molecules (3.6.1, 3.6.2), in glycoproteins (4.4) and glycolipids (4.4) and as structural elements in cell walls of bacteria (3.10.1), plants (3.4) and in the exoskeleton of arthropods (3.1). They are the most abundant chemical group in the biosphere.

1.2.1 Structure and Classification

The simplest carbohydrates are the trioses (C_3 compounds) glyceraldehyde (an aldose) and dihydroxyacetone (glycerone, a ketose). Larger molecules are tetroses (C_4), pentoses (C_5), hexoses (C_6), heptoses (C_7) etc.; the C_5 and C_6 molecules are most common.

Glyceraldehyde is the smallest aldose with an asymmetric C-atom (chirality center). Therefore there are two stereoisomers (enantiomers), which cause right and left rotation of polarized light. By the Fischer convention, they are named D- and L-form, respectively. For details, see organic chemistry textbooks. Tetroses and larger carbohydrate monomers are classified (by comparison of the asymmetric center most distant to the aldehyde or keto group with D- or L-glyceraldehyde) as the D- and L-series of enantiomers (Fig. 1.2-1). With n-carbon aldoses, a total of 2^{n-2} stereoisomers exist, and with n-carbon ketoses there are 2^{n-3} stereoisomers. Epimers are stereoisomers, which differ in configuration at only one asymmetric C-atom. Most physiological sugars are of the D-configuration.

Aldopentoses, aldohexoses and ketohexoses (and higher sugars) can form cyclic structures (hemiacetals and hemiketals) by intramolecular reaction of their aldehyde or keto groups respectively with an alcohol group. This results in pyranoses (6-membered rings) and furanoses (5-membered rings, Fig. 1.2-2). In equilibrium, the cyclic structure is more prevalent as compared to the open structure. The ring closure produces another asymmetric C-atom; the respective stereoisomers are named anomers (α- and β-forms).

The nonplanar pyranose rings can assume either boat (in 2 variants) or chair conformation. The substituents extend either parallel to the perpendicular axis (axial, in Fig. 1.2-3 printed in red) or at almost right angles to it (equatorial, printed in green). The preferred

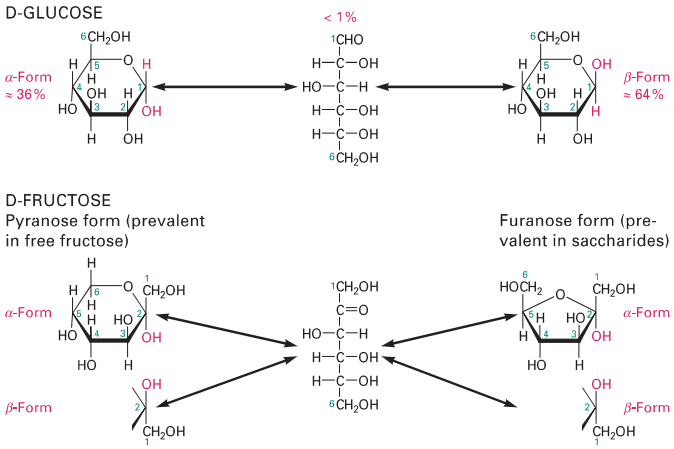


Figure 1.2-2. Ring Closure of Carbohydrates

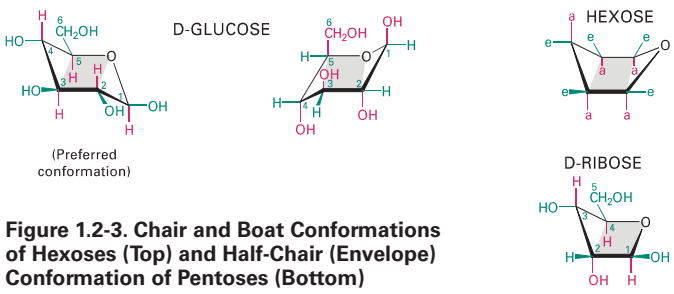


Figure 1.2-3. Chair and Boat Conformations of Hexoses (Top) and Half-Chair (Envelope) Conformation of Pentoses (Bottom)

conformation depends on spatial interference or other interactions of the substituents.

Although the bond angles of a furanose ring would permit an almost planar structure, the interference of substituents with each other causes a slight bending (puckering), e.g., to a half-chair (= envelope) structure in nucleotides and nucleic acids (Fig. 1.2-3).

The linear form of carbohydrates is usually shown as Fischer projection (ligands drawn horizontally are in front of the plane, ligands drawn vertically are behind the plane, e.g., in Fig. 1.2-1). The ring form is either drawn as Haworth formula (Fig. 1.2-2, disregarding the bent ring structure) or as boat/chair formula.

1.2.2 Glycosidic Bonds (Fig. 1.2-4)

If the hemiacetal or hemiketal hydroxyl of a sugar is condensed with an alcoholic hydroxyl of another sugar molecule, a glycosidic bond is formed and water is eliminated. Since this reaction between free sugars is endergonic ($\Delta G'_0 = 16$ kJ/mol), the sugars usually have to be activated as nucleotide derivatives (3.1.2.2) in order to be noticeably converted. Depending on the configuration at the hemiacetal/hemiketal hydroxyl (1.2.1), either α- or β-glycosides are formed. Sugar derivatives, which contain a hemiacetal or a hemiketal group (e.g., uronic acids) are also able to form glycosidic bonds.

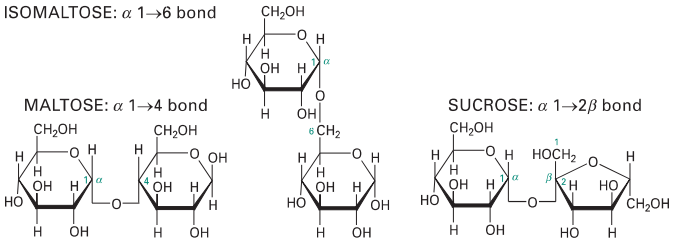


Figure 1.2-4. Examples of Glycosidic Bonds

Since sugar molecules contain several alcoholic groups, various types of bonds are possible. Frequently, 1 → 4 or 1 → 6 bonds occur. With oligo- or polysaccharides, both linear and branched structures are found. Bond formation may also take place with alcoholic, phenolic or other groups of non-sugar molecules (aglycons).

Literature:

Organic chemistry textbooks.

1.3 Amino Acid Chemistry and Structure

All amino acids present in proteins carry a carboxyl- and an amino group, hydrogen and variable side chains (R) at a single (α-)carbon atom. Thus, this C_α-atom is asymmetric (compare 1.2.1), with the exception of glycine, where R = H. Almost all of the proteinogenic amino acids occurring in nature are of the L-configuration. (The 'L' is assigned by comparison with L- and D-glyceraldehyde, which are taken as standards, Fig. 1.3-1). A number of D-amino acids are found in bacterial envelopes (3.10) and in some antibiotics (3.10).

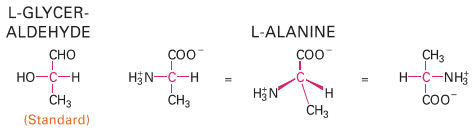


Figure 1.3-1. Asymmetric Center of Amino Acids

Unless otherwise stated, all amino acids discussed in the following sections are of the L-configuration.

Chains of amino acids form proteins and peptides. As enzymes, regulatory, mobility and structural compounds, they are the central components in all living beings. Therefore they are the topic of most of this book. Protein synthesis is described in Chapter 4. Their structure is discussed in Section 2.3, which also gives a short listing of their functions.

a) Non-polar, aliphatic amino acids. The non-polar side chains undergo hydrophobic interactions in protein structures. While the small glycine molecule allows high flexibility, the bulky proline confers enhanced rigidity to the structures.

GLYCINE (Gly, G) L-ALANINE (Ala, A) L-VALINE (Val, V) L-LEUCINE (Leu, L) L-ISOLEUCINE (Ile, I) L-PROLINE (Pro, P)

b) Polar, uncharged residues R. These functional groups are hydrophilic and can form hydrogen bonds with water or other polar compounds. Cysteine can easily be oxidized, resulting in intra- or intermolecular interconnections by disulfide bonds.

L-SERINE (Ser, S) L-THREONINE (Thr, T) L-CYSTEINE (Cys, C) L-METHIONINE (Met, M) L-ASPARAGINE (Asn, A) L-GLUTAMINE (Gln, Q)

c) Aromatic residues R. The aromatic side chains are hydrophobic, while the hydroxyl group of tyrosine and the ring nitrogen of tryptophan form hydrogen bonds, which often play a role in enzyme catalysis.

L-PHENYLALANINE (Phe, F) L-TYROSINE (Tyr, Y) L-TRYPTOPHAN (Trp, W)

d) Positively charged side chains R. The charged groups contribute in many cases to catalytic mechanisms and also influence the protein structure.

L-LYSINE (Lys, K) L-ARGININE (Arg, R) L-HISTIDINE (His, H)

e) Negatively charged side chains R. The charged groups contribute in many cases to catalytic mechanisms and are also of influence to the protein structure.

L-ASPARATE (Asp, D) L-GLUTAMATE (Glu, E)

L-PYRROLYSINE (Pyl, O) L-SELENOCYSTEINE (Sec, U)

f) Special amino acids using unusual genetic codes and found only in a small number of proteins (3.10.6.2, 4.1.3.4).

Figure 1.3-2. Amino Acids With Their 3- and 1-Letter Codes

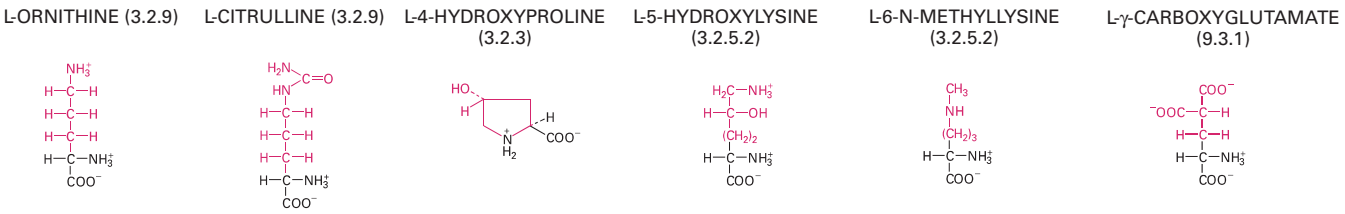


Figure 1.3-3. Some Nonstandard Amino acids

1.3.1 Structure and Classification

The individual properties of the amino acids are determined by the side chain R. This is also the criterion for amino acid classification.

There are 20 standard (classical) amino acids, which are incorporated as such into proteins, employing their own codons (4.1, 4.2). These amino acids are shown in Figure 1.3-2. Two additional amino acids, selenocysteine and pyrrolysine, are also incorporated directly by an unusual decoding procedure of mRNA (4.1). Nonstandard amino acids are produced by metabolic conversions of free amino acids (e.g., ornithine and citrulline) or by posttranslational modification of amino acids in proteins (e.g., by hydroxylation, methylation or carboxylation). Examples are given in Figure 1.3-3.

At about neutral pH, the free amino acids are 'Zwitterion' dipols with charged carboxylate (dissociation constant pK_1 = 1.82 … 2.35) and amino groups (pK_2 = 8.70 … 10.70). In seven cases, the side chains R also contain charged groups. Only the pK_α of histidine (3.2.8) is in the physiological range. In Figure 1.3-2 and 1.3-3, the charged molecules are shown, while in the rest of the book, un-ionized forms are presented for reasons of simplicity.

1.3.2 Peptide Bonds (Fig. 1.3-1)

Proteins and peptides are linear chains of amino acids connected by peptide bonds between their α-amino and carboxylate groups. Since the formation of these bonds is endergonic, the reactants have to be activated as tRNA derivatives. Details are described in 4.1.3.

The peptide bonds are rigid and planar: The carboxylate-O and the amino-H are in *trans* conformation, the C–N bond shows partially double bond characteristics. Only peptide bonds followed by proline or hydroxyproline can alternatively be *cis* (6 … 10 %). To some extent, both bonds in the backbone of the peptide chain extending from C_α can perform rotational movements (although there are still constraints on most conformations, which are shown in Ramachandran diagrams). Flexibility and constraints play a major role in the proper folding of the proteins (1.3.1).

Proteins and peptides carry charged amino- (N-) and carboxy- (C-) termini. Additional charges are contributed by the side chains. This allows analytical separation by electrophoresis. It has to be considered, however, that the pK_α of amino acids in peptides differ from those in free amino acids due to the effects of neighboring groups.

Literature:

Meister, A. *Biochemistry of the Amino Acids*. 2 Vols. Waltham, (MA):Academic Press; 1965.

Ramachandran, G.N., Sasisekharan, V. Adv. Prot. Chem. 1968;23:326–367.

Rose, G.D. *et al.* Adv. Prot. Chem. 1985;37:1–109.

Organic chemistry textbooks.

1.4 Lipid Chemistry and Structure

The common properties of lipids are their hydrophobic character and their solubility in organic solvents. Otherwise, they belong to different chemical classes. The biochemistry of most of them is described in Chapter 3.4, some other lipids are discussed in their metabolic context elsewhere (see cross-references below).

1.4.1 Fatty acids (Table 1.4-1, Fig. 1.4-1)

Fatty acids are characterized by a carboxylic group with a hydrocarbon 'tail'. The higher fatty acids are practically insoluble in water and show typical lipid properties. They serve in esterified form as triacylglycerols for energy storage or are, as glycerophospholipids, part of cellular membranes. In contrast, the short-chain fatty acids are water soluble. They act as intermediates of metabolism and are discussed in the respective chapters.

Higher fatty acids can also enter an amide bond (e.g., in ceramides). Some are precusors of other compounds (e.g., of prostaglandins, 7.4.8). Almost none of them occur in free form.

The predominant fatty acids in higher plants and animals have an even number of C atoms in the range of C_{14} … C_{20} and are unbranched. Usually, more than half of all fatty acids are unsaturated. Monounsaturated fatty acids mostly contain a *cis*-double bond between C-9 and C-10. Often additional double bonds exist towards the methyl terminus, usually with two saturated bonds in between (polyunsaturated fatty acids). Some of them cannot be synthesized in animals and have to be supplied by food intake (essential fatty acids). The notation of fatty acids is (number of C atoms) : (number of double bonds), e.g., for linoleic acid 18:2. The location of the double bonds is given as, e.g., $\Delta^{9,12}$.

Polyunsaturated fatty acids are not usually present in bacteria, but *cis*- and *trans*-monounsaturated, hydroxylated and branched fatty acids exist in many species.

Table 1.4-1. Higher Fatty Acids Frequently Occurring in Nature

Number of C atoms	Saturated	Unsaturated, Number of Double Bonds	E = Essential Fatty Acid for Humans
14	myristic acid		
16	palmitic acid		
18	stearic acid	1: oleic acid (Δ^9)	
		2: linoleic acid ($\Delta^{9,12}$)	E
		3: α-linolenic acid ($\Delta^{9,12,15}$)	E
20	arachidic acid	3: dihomo-γ-linolenic acid ($\Delta^{5,8,11}$)	
		4: arachidonic acid (AA, $\Delta^{5,8,11,14}$)	(E)[1]
		5: eicosapentaeonoic acid (EPA, $\Delta^{5,8,11,14,17}$)	
24	lignoceric acid	1: nervonic acid (Δ^{15})	

[1] can be synthesized from the essential fatty acid, linoleic acid.

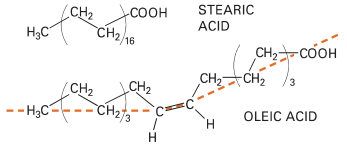


Figure 1.4-1. Structure of Saturated and Unsaturated Fatty Acids (18:0 and 18:1, showing the bend).

While saturated fatty acids tend to assume an extended shape, unsaturated fatty acids show 30° bends at their double bonds (Fig. 1.4-1). This reduces van der Waals interactions between neighboring molecules and lowers the melting point (see organic chemistry textbooks):

18:0 (70°) 18:1 (13°) 18:2 (–9°) 18:3 (–17°)

1.4.2 Acylglycerols and Derivatives (Fig. 1.4-2)

A major proportion of lipids occurring in plants and animals are triesters of glycerol (3.4.2) with higher fatty acids (triacylglycerols = triglycerides = neutral fat). In most of them, the fatty acids are different. Their type and the degree of their unsaturation determine the melting point.

Fats are solid and oils are liquid at room temperature. They are without influence on the osmotic situation in the aqueous phase due to their insolubility and do not bind water as, e.g., glycogen does. Thus, these compounds constitute an effective, convenient storage form of energy (ca.10 kg in adult humans).

Their degree of oxidation is lower than that of carbohydrates or proteins, therefore they provide higher energy during combustion: triolein yields 39.7 kJ/g. This is more than twice the value for anhydrous carbohydrates (17.5 kJ/g) or proteins (18.6 kJ/g) and about six times the energy gained from degradation of these alternative compounds in their physiological state due to their water content.

Triacylglycerols do not contain any hydrophilic groups. If, however, only one or two of the hydroxyl groups of glycerol are esterified (mono- or diacylglycerols), the remaining polar hydroxyl groups allow the formation of ordered structures at water-lipid interfaces and of lipid bilayers (1.4.8). Therefore they can act as emulsifiers, e.g., during lipid resorption from the intestine.

The remaining hydroxyl groups of mono- and diacylglycerols can also carry sugar residues. These so-called glycoglycerolipids are constituents of bacterial cell envelopes (3.10), thylakoid membranes in plants and myelin sheaths of neurons in animals. They are discussed in 3.4.

1.4.3 Waxes (Table 1.4-2, Fig. 1.4-2)

Waxes are esters of higher fatty acids with long-chain primary alcohols (wax alcohols) or sterols (Section 3.5), which are usually solid at room temperature.

They are more resistant than triacylglycerols towards oxidation, heat and hydrolysis (saponification). Frequently, they serve as protective layers, e.g., on leaves and fruits of plants or on skin, feathers and furs of animals (as secretions of specialized glands). Bees' honeycombs are also formed of waxes. In many marine animals they are the main component of lipids (for regulation of flotation and for energy storage). Fossil waxes occur in lignite and bitumen.

Table 1.4-2. Common Components of Waxes

Alcohol (primary, saturated)	Fatty Acid (saturated)
Cetyl alcohol (C_{16})	Lauric acid (C_{12})
Carnaubyl alcohol (C_{24})	Myristic Acid (C_{14})
Ceryl alcohol (C_{26})	Palmitic acid (C_{16})
Myricyl alcohol (C_{30})	Lignoceric Acid (C_{24})
	Cerotic Acid (C_{26})
	Montanic Acid (C_{28})
	Melissic Acid (C_{30})

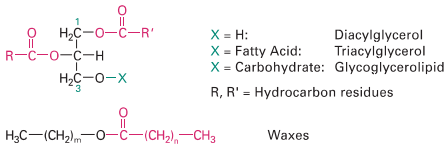
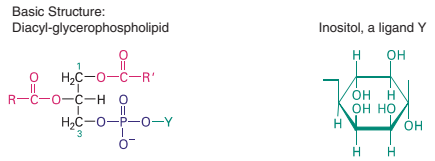


Figure 1.4-2. Structure of Acylglycerols, Glycoglycerolipids and Waxes

1.4.4 Glycerophospholipids (Phosphoglycerides, Fig. 1.4-3)

In contrast to triacylglycerols, in glycerophospholipids only two of the hydroxyl groups of glycerol are esterified with long chain fatty acids, while the group at the 3-position (according to the *sn*-numbering system) forms an ester with phosphoric acid.

All glycerophospholipids have an asymmetric C-atom in the 2-position, they occur in nature in the L-form. Most common are saturated fatty acids (C_{16} or C_{18}) at the 1- and unsaturated ones (C_{16} … C_{20}) at the 2-position. Removal of one fatty acid yields lysoglycerophospholipids.



Class	Y =	Formula of Y (F = Fatty acid)
Phosphatidic acid	H	–H
Phosphatidylethanolamine	ethanolamine	$–CH_2–CH_2–NH_2$
Phosphatidylcholine (lecithin)	choline	$–CH_2–CH_2–N(CH_3)_3^+$
Phosphatidylserine	serine	$–CH_2–CH(NH_2)–COOH$
Phosphatidylinositol	inositol	see above
Diphosphatidylglycerol (cardiolipin)	glycerol-phosphatidic acid	$–CH_2–CHOH–CH_2O–(PO_2)–$ $–O–CH_2–CHOF–CH_2O–F$

Figure 1.4-3. Classes of Glycerophospholipids

If the 3-position of glycerol carries only phosphoric acid, the compound is named phosphatidic acid. However, in most cases the phosphate group is diesterified. This extra residue ('head group', Y in Fig. 1.4-3) determines the class of the compound. These compounds are more polar than mono- or diacylglycerols and form the major part of biological membranes (1.4.8).

1.4.5 Plasmalogens (Fig. 1.4-4)

This group of compounds is related to diacylglycerophospholipids (1.4.4). Also, the head groups (Y) are similar. However, the 1-position of glycerol is not esterified, but carries an α,β-unsaturated alcohol in an ether linkage. They are major components of the CNS, brain (>10 %), heart and skeletal muscles, but little is known about their physiological role.

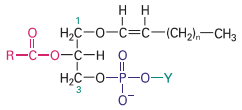
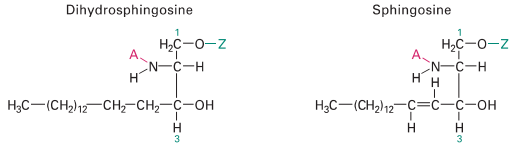


Figure 1.4-4. Structure of Plasmalogens

1.4.6 Sphingolipids (Fig. 1.4-5)

Sphingolipids are important membrane components. They are derivatives of the aminoalcohols dihydrosphingosine (C_{18}), sphingosine (C_{18} with a *trans* double bond) or their C_{16}, C_{17}, C_{19} and C_{20} homologues.

Ceramides are N-acylated sphingosines. If the hydroxyl group at C-1 is esterified with phosphocholine, phosphoethanolamine etc., sphingomyelins (sphingophospholipids) are obtained. If, alternatively, the hydroxyl group is glycosylated, glycosphingolipids (cerebrosides) result. This latter group of compounds is described in 4.4.2-3.



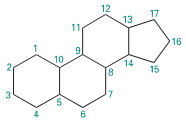
Class	A (= N-Substituent)	Z	Formula of Z
Sphingosine	H	H	–H
Ceramides	higher fatty acid	H	–H
Sphingomyelins	higher fatty acid	Phosphoethanolamine or phosphocholine	$–PO_2–O-CH_2–CH_2–OH$ $–PO_2–O–CH_2–CH_2–N(CH_3)_3^+$
Glycosphingolipids	higher fatty acid	carbohydrate chains	-carbohydrates

Figure 1.4-5. Basic Structure of Sphingolipids

1.4.7 Steroids

Steroids are derivatives of the hydrocarbon cyclopentanoperhydrophenanthrene (Fig. 1.4-6).

Figure 1.4-6. Structure of Cyclopentanoperhydrophenanthrene



Biologically important steroids carry many substituents: generally there is a hydroxy or oxo group at C-3. In addition, several methyl, hydroxy and oxo, in some cases also carboxy, groups are found. In many cases, there is a larger residue bound to C-17. Frequently, some double bonds are present. In a few cases, ring A is aromatic. Substituents below the ring system are designated α and above the ring system β (see Fig. 3.5.1-5).

Steroids are membrane components and participants as well as regulators of metabolism. A detailed description is given in Section 3.5.

1.4.8 Lipoproteins

The major function of lipoproteins is the transport of lipids. They contain non-polar lipids (triacylglycerols, cholesterol esters) in their core, surrounded by a layer of polar compounds (glycerophospholipids, cholesterol, proteins, Fig. 6.2-1). This group of compounds is discussed in context with their transport function in 6.2.

1.5 Physico-Chemical Aspects of Biochemical Processess

Some readers may be less inclined to deal with a fairly large number of mathematical formulas. However, formulas are necessary to describe biochemical processes quantitatively. Considering this, the mathematical part of this book has been concentrated into this section, while usually other chapters refer to it.

Only the most important equations required for discussion of biochemical reactions are presented. In order to facilitate their use, companion equations are given, which show the numerical values of the factors and the dimensions of the terms. For derivation of the equations, refer to physical chemistry textbooks. The units and constants used in the following paragraphs are listed in Table 1.5-1.

1.5.1 Energetics of Chemical Reactions

To each component of a system, an amount of free energy G is assigned, which is composed of the enthalpy H (internal energy + pressure * volume) and of the entropy S (measure of disorder). While the absolute values are not of importance, the change of G (ΔG) is decisive for chemical reactions:

$$\Delta G = \Delta H - T * \Delta S \qquad [1.5\text{-}1]$$

or

$$\Delta G\ [kJ * mol^{-1}] = \Delta H\ [kJ * mol^{-1}] - T\ [K] * \Delta S\ [kJ*mol^{-1} * K^{-1}]. \qquad [1.5\text{-}1a]$$

A reaction proceeds spontaneously only if ΔG is negative.

In biochemistry, ΔG of reactions are usually listed as $\Delta G'_0$, which is obtained at standard conditions of 298 K (25 °C), pH 7.0 and a reactant concentration of 1 mol/l each except for water, where the normal concentration of 55.55 mol/l and gases, where a pressure of 101.3 kPa (= 1 atm) are taken as unity and thus do not appear in the formula.

If the reactant concentrations (henceforth written as [X]) of a reaction A + B + … = Z + Y + … differ from 1 mol/l each, ΔG can be calculated by:

$$\Delta G = \Delta G'_0 + R * T * 2.303 * \log \frac{[Z] * [Y] * \ldots^{\text{(end products)}}}{[A] * [B] * \ldots^{\text{(starting comp.)}}} \qquad [1.5\text{-}2]$$

or

$$\Delta G\ [kJ * mol^{-1}] = \Delta G'_0 + 0.00831 * T * 2.303 * \log \frac{[Z] * [Y] * \ldots}{[A] * [B] * \ldots} \qquad [1.5\text{-}2a]$$

Reaction sequences can be calculated by addition of ΔG's of the individual reactions.

A reaction is at equilibrium if ΔG = 0. Then the equilibrium constant

$$K = \frac{[Z] * [Y] * \ldots^{\text{(end products)}}}{[A] * [B] * \ldots^{\text{(starting comp.)}}} \qquad [1.5\text{-}3]$$

can be calculated as follows:

$$\Delta G'_0 = -R * T * 2.303 * \log K; \qquad K = 10^{(-\Delta G / R * T * 2.303)} \qquad [1.5\text{-}4]$$

or

$$\Delta G'_0\ [kJ * mol^{-1}] = -0.00831 * T * 2.303 * \log K; \qquad K = 10^{(-\Delta G / 0.00831 * T * 2.303)} \qquad [1.5\text{-}4a]$$

Table 1.5-1. Measures and Constants (Selection)

Measure	Unit	Equivalents / Value of Constants	Equivalents in SI Basic Units
Length	meter (m)	1 mm = 10^{-3} m, 1 µm = 10^{-6} m, 1 nm = 10^{-9} m, 1 Å (Ångström) = 10^{-10} m	SI basic unit
Volume	cubic meter (m^3)	1 l (liter) = 10^{-3} m^3; 1 ml = 10^{-3} l, 1 µl = 10^{-6} l	Derived SI unit
Mass	kilogram (kg)	1 g (gram) = 10^{-3} kg; 1 mg = 10^{-3} g, 1 µg = 10^{-6} g	SI basic unit
Time (t)	second (s)	1 ms = 10^{-3} s, 1 µs = 10^{-6} s, 1 ns = 10^{-9} s, 1 ps = 10^{-12} s	SI basic unit
Temperature (T)	Kelvin (K)	0 K = –273.16 °C	SI basic unit
Quantity of matter	Mol	1 mol = 6.0221 * 10^{23} [molecules or ions] 1 mmol = 10^{-3} mol, 1 µmol = 10^{-6} mol, 1 nmol = 10^{-9} mol, 1 pmol = 10^{-12} mol This unit is also applied to photons. 1 Einstein = 1 mol photons.	SI basic unit
Electric current	Ampere (A)		SI basic unit
Force	Newton (N)	1 N = 1 [m * kg * s^{-2}]	1 N = 1 [m * kg * sec^{-2}]
Pressure	Pascal (Pa)	1 Pa = 1 N * m^{-2}, 1 kilopascal (kPa) = 10^3 Pa, 1 atm = 101.325 kPa	1 Pa = 1 [m^{-1} * kg * sec^{-2}]
Energy	Joule (J)	1 J = 1 [N * m], 1 kilojoule (kJ) = 1000 J 1 cal (calorie) = 4.181 J, 1 kcal = 4.181 kJ (Non-SI unit)	1 J = 1 [m^2 * kg * sec^{-2}]
Electric charge	Coulomb (C)	1 C = 6.241 * 10^{18} electron charges	1 C = 1 [A * sec]
Electric potential	Volt (V)	1 V = 1 [J * C^{-1}], 1 mV = 10^{-3} V	1 V = 1 [m^2 * kg * sec^{-3} * A^{-1}]
Constants	**Abbreviation**	**Equivalents / Value of Constants**	**Equivalents in SI Basic Units**
Avogadro's number	N	N = 6.0220 * 10^{23} [mol^{-1}] (see 'quantity of matter', above)	N = 6.0220 * 10^{23} [mol^{-1}]
Boltzmann's constant	k_B	k_B = 1.3807 * 10^{-23} [J * K^{-1}]	k_B = 1.3807 * 10^{-23} [m^2 * kg * sec^{-2} * K^{-1}]
Molar gas constant	R	R = N * k_B = 8.31441 [J * mol^{-1} * K^{-1}]	R = 8.31441 [m^2 * kg * sec^{-2} * mol^{-1} * K^{-1}]
Faraday's constant	F	F = 1 N electron charges = 96 484.5 [C * mol^{-1}] = 96 484.5 [J * V^{-1} * mol^{-1}]	F = 96 484.5 [A * sec * mol^{-1}]
Planck's constant	h	h = 6.6262 * 10^{-34} [J * sec]	h = 6.6262 * 10^{-34} [m^2 * kg * sec^{-1}]

In calculations in this book, usually l, g, kJ and mV are used. Since the constants have then to be expressed in these units, their numerical value changes by the factor 10^{-3} or 10^{-6} in the respective formulas.

Enzymes cannot shift the equilibrium, they only increase the reaction velocity. The kinetics of enzyme catalyzed reactions are discussed in 1.5.4.

1.5.2 Redox Reactions

Redox reactions are reactions where one compound is reduced (electron acceptor A) while its reaction partner is oxidized (electron donor B) by transfer of n electrons:

$$A_{ox}^{n+} + B_{red} = A_{red} + B_{ox}^{n+}.$$

The change of free energy during such a reaction is described by a formula, which is analogous to Eq. [1.5-2]:

$$\Delta G = \Delta G_0' + R * T * 2.303 * \log \frac{[A_{red}] * [B_{ox}^{n+}]\ \text{(end products)}}{[A_{ox}^{n+}] * [B_{red}]\ \text{(starting comp.)}} \qquad [1.5\text{-}5]$$

or

$$\Delta G\ [\text{kJ} * \text{mol}^{-1}] = \Delta G_0' + 0.00831 * T * 2.303 * \log \frac{[A_{red}] * [B_{ox}^{n+}]}{[A_{ox}^{n+}] * [B_{red}]} \qquad [1.5\text{-}5a]$$

w expresses the work gained by transferring n mol charges (= n Faraday, F) across a potential difference of $\Delta E = E_{end} - E_{begin}$

$$w = -n * F * \Delta E. \qquad [1.5\text{-}6]$$

Since a positive amount of work diminishes the free energy of the system

$$w = -n * F * \Delta E = -\Delta G \qquad [1.5\text{-}6a]$$

or

$$\Delta G\ [\text{kJ} * \text{mol}^{-1}] = n * 0.0965 * \Delta E\ [\text{mV}], \qquad [1.5\text{-}6b]$$

equation [1.5-5] can also be written as:

$$\Delta E = \Delta E_0' + \frac{R * T}{n * F} * 2.303 * \log \frac{[A_{red}] * [B_{ox}^{n+}]\ \text{(end products)}}{[A_{ox}^{n+}] * [B_{red}]\ \text{(starting comp.)}} \qquad [1.5\text{-}7]$$

or

$$\Delta E\ [\text{mV}] = \Delta E_0' + \frac{0.00831 * T}{n * 0.0965} * 2.303 * \log \frac{[A_{red}] * [B_{ox}^{n+}]}{[A_{ox}^{n+}] * [B_{red}]} \qquad [1.5\text{-}7a]$$

$\Delta E_0'$ is the difference of the redox potentials of this reaction (or the electromotive force across membranes, 1.5.3) under biochemical standard conditions (298 K = 25°C, pH 7.0 and a reactant concentration of 1 mol/l each). Only water, which is present in a concentration of 55.55 mol/l and gases, with a pressure of 1 atm are taken as unity.

Redox potentials: The reaction can be divided into two half reactions (e^- = electrons):

$$A_{red} = A_{ox}^{n+} + n\ e^- \quad \text{and} \quad B_{red} = B_{ox}^{n+} + n\ e^-$$

The zero value of the redox potential is by convention assigned to the potential of the half reaction $2\ H^+ + 2\ e^- = H_2$ at a platinum electrode at pH = 0, 298 K (25°C) and a hydrogen pressure of 101.3 kPa (= 1 atm). Thus, under the standard conditions used in biochemistry (pH = 7.0), $E_0'\ (2H^+/H_2) = -410$ mV.

Correspondingly, the half reactions can be expressed as:

$$E_A = (E_0')_A + \frac{R * T}{n * F} * 2.303 * \log \frac{[A_{ox}^{n+}]}{[A_{red}]} \qquad [1.5\text{-}8]$$

or

$$E_A\ [\text{mV}] = (E_0')_A + \frac{0.00831 * T}{n * 0.0965} * 2.303 * \log \frac{[A_{ox}^{n+}]}{[A_{red}]} \qquad [1.5\text{-}8a]$$

and analogously for B.

Various redox potentials can be combined this way: $\Delta E = E_B - E_A$ (A being the electron acceptor and B being the electron donor). The reactions proceed spontaneously only if ΔE is negative, i.e., when the potential changes to a more negative value.

Redox potentials are usually plotted with the minus values on top. A spontaneous reaction proceeds in such a plot from top to bottom (e.g., Fig. 3.12-6).

In the literature, the definition of ΔE is not uniform. In a number of textbooks it is defined in opposite order to the above: $\Delta E = E_{begin} - E_{end}$. Therefore, ΔE and ΔE_0 have to be replaced by $-\Delta E$ and $-\Delta E_0'$, respectively. This affects Eqs. [1.5-6] … [1.5-8a] and has to be considered when making comparisons.

1.5.3 Transport Through Membranes

Uncharged molecules: If an uncharged compound A is present on both sides of a permeable membrane in different concentrations, its passage through the membrane is accompanied by a change of free energy. In biochemistry, this situation occurs mostly at cellular membranes (or membranes of organelles). For import into cells, the following equation applies:

$$\Delta G = R * T * 2.303 * \log \frac{[A_{inside}]}{[A_{outside}]} \qquad [1.5\text{-}9]$$

or

$$\Delta G\ [\text{kJ} * \text{mol}^{-1}] = 0.00831 * T * 2.303 * \log \frac{[A_{inside}]}{[A_{outside}]} \qquad [1.5\text{-}9a]$$

Thus, the transport occurs spontaneously only at negative ΔG, (when $[A_{inside}] < [A_{outside}]$), i.e, from higher to lower concentrations.

Correspondingly, for export from cells, the quotient is reversed

$$\Delta G = R * T * 2.303 * \log \frac{[A_{outside}]}{[A_{inside}]} \qquad [1.5\text{-}9b]$$

Charged molecules: The situation is more complicated if there is a potential difference $\Delta\Psi$ across the membrane (e.g., by non-penetrable ions)

$$\Delta\Psi = \Psi_{inside} - \Psi_{outside} \qquad [1.5\text{-}10]$$

and the compounds passing through the membrane carry Z positive charges/molecule (or −Z negative charges/molecule). The contribution of the charges to ΔG (with the prefix of Z corresponding to the + or − charge of the ions) is expressed by:

$$\Delta G_{charge\ transport} = Z * F * \Delta\Psi \qquad [1.5\text{-}11]$$

or

$$\Delta G_{charge\ transport}\ [\text{kJ} * \text{mol}^{-1}] = Z * 0.0965 * \Delta\Psi\ [\text{mV}]. \qquad [[1.5\text{-}11a]$$

Thus, for an import process, Eq. [1.5-9] and Eq. [1.5-11] have to be combined:

$$\Delta G = R * T * 2.303 * \log \frac{[A_{inside}]}{[A_{outside}]} + Z * F * \Delta\Psi \qquad [1.5\text{-}12]$$

or

$$\Delta G\ [\text{kJ} * \text{mol}^{-1}] = 0.00831 * T * 2.303 * \log \frac{[A_{ins.}]}{[A_{outs.}]} + Z * 0.0965 * \Delta\Psi\ [\text{mV}]. \qquad [1.5\text{-}12a]$$

Correspondingly, for an export process,

$$\Delta G = R * T * 2.303 * \log \frac{[A_{outside}]}{[A_{inside}]} - Z * F * \Delta\Psi \qquad [1.5\text{-}12b]$$

The prefix of the last term in this equation is the opposite one of Eq. [1.5-12], since the membrane potential (Eq. 1.5-10) has the opposite effect on the energy situation.

An equilibrium exists if $\Delta G = 0$. Then the equilibrium potential $\Delta\Psi_0$ [mV] can be obtained by the Nernst equation:

$$\Delta\Psi_0 = -\frac{R * T}{Z * F} * 2.303 * \log \frac{[A_{inside}]}{[A_{outside}]} \qquad [1.5\text{-}13]$$

or

$$\Delta\Psi_0\ [\text{mV}] = -\frac{0.00831 * T}{Z * 0.0965} * 2.303 * \log \frac{[A_{inside}]}{[A_{outside}]} \qquad [1.5\text{-}13a]$$

An extension of this formula to the equilibrium potential of several ions is the Goldman equation (see 7.2.1).

Literature:

Physical chemistry textbooks.

1.5.4 Enzyme Kinetics

The biochemical base of enzyme catalysis is discussed in 2.4. In the following, the mathematical treatment of the kinetics is given in some more detail.

Velocity of reactions: The reaction rate v for conversion of a single compound A → product(s) (first order reaction) is proportional to the concentration of this compound [A], while for a two-compound reaction A + B → product(s) (second order reaction) it depends on the number of contacts and thus on the concentration of both components (Eq. [1.5-14] and Eq. [1.5-15]). The proportionality factor k is termed rate constant.

Eq. [1.5-15] can also be applied for the formation of a complex and Eq. [1.5-14] for the decomposition of this complex. This includes substrate-enzyme complexes (see below), ligand-receptor complexes (7.1-2), antigen-antibody complexes (8.1.4) etc.

$$v = -\frac{d[A]}{dt} = k * [A] \qquad [1.5\text{-}14]$$

$$v = -\frac{d[A]}{dt} = -\frac{d[B]}{dt} = k * [A] * [B] \qquad [1.5\text{-}15]$$

Enzyme catalyzed one-substrate reaction: The theory of the enzyme-catalyzed conversion of a single reactant (the substrate, S) is based on the assumption that the enzyme (the catalyst, E) and this substrate form a complex (ES) by a reversible reaction. This step is kinetically treated like a two-compound reaction (rate constants k_1 and k_{-1} for formation and decomposition, respectively). The complex is then converted into the product (P) with the rate constant k_2. The conversion into P is considered to be irreversible at the beginning of the reaction, when practically no product is present.

$$E + S \underset{k_{-1}}{\overset{k_1}{\leftrightarrow}} ES \xrightarrow{k_2} E + P \qquad [1.5\text{-}16]$$

Therefore, for the formation of the enzyme-substrate complex, Eq. [1.5-15] has to be applied, while for its decomposition into its components, as well as for its conversion to the products, Eq. [1.5-14] is valid. There is actually an intermediate step ES → EP before the product is released. Its rate constant is not treated as a separate entity in most discussions of kinetic behavior, but is combined with the dissociation step to k_2. This is also done in the following considerations.

Usually, the substrate is in large excess over the enzyme. In this case, after a short 'transient phase', [ES] can be considered to be sufficiently constant (steady-state assumption). Disregarding the reverse reaction by using the situation immediately after the transient phase (see above) one obtains

$$\frac{d[ES]}{dt} = 0 = k_1 * [E] * [S] - k_{-1} * [ES] - k_2 * [ES] \qquad [1.5\text{-}17]$$

If one assumes that the rate determining process is the reaction ES → E + P, the initial reaction rate v_0 can be written as a function of [ES], which is analogous to Eq. [1.5-14]

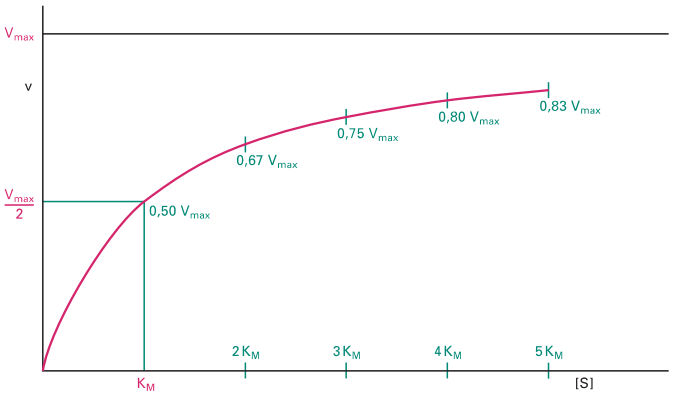


Figure 1.5-1. Reaction Velocity of an Enzyme Catalyzed Reaction
The velocity at [S] = n * K_M is shown.

$$v_0 = k_2 * [ES]. \qquad [1.5\text{-}14a]$$

By using a term for the total concentration of enzyme $[E_t] = [E] + [ES]$, by expressing the maximum reaction rate V_{max}, which is obtained when all of the enzyme is saturated with substrate ([ES] = $[E_t]$) as

$$V_{max} = k_2 * [E_t], \qquad [1.5\text{-}14b]$$

and by introducing the Michaelis constant K_M

$$K_M = \frac{k_{-1} + k_2}{k_1}, \qquad [1.5\text{-}18]$$

one obtains the so-called Michaelis-Menten equation

$$v_0 = \frac{V_{max} * [S]}{K_M + [S]}, \qquad [1.5\text{-}19]$$

which shows the dependency of the reaction rate on the substrate concentration (first-order reaction). The plot of reaction rate vs. substrate concentration is a rectangular hyperbola (Fig. 1.5-1).

These formulas describe only the forward reaction. If the reverse reaction is included, the equivalent to Eq. [1.5-19] is

$$v = \frac{\frac{(V_{max})_f * [S]}{(K_M)_f} - \frac{(V_{max})_r * [P]}{(K_M)_r}}{1 + \frac{[S]}{(K_M)_f} + \frac{[P]}{(K_M)_r}} \qquad [1.5\text{-}20]$$

where $(V_{max})_f$ and $(K_M)_f$ are identical to V_{max} and K_M in Eq. [1.5-19], while the terms $(V_{max})_r$ and $(K_M)_r = (k_{-1} + k_2)/k_{-2}$ are formed analogously for the reverse reaction.

Michaelis constant: As can be derived from Eq. [1.5-19], the Michaelis constant K_M equals the substrate concentration at half the maximal reaction rate. Most of them are in the range of 10^{-5}…10^{-1} mol/l (Fig. 10.3-3).

Instead of obtaining this value from a plot according to Figure 1.5-l, it is more convenient to use the reciprocal of the Michaelis-Menten equation, which yields a linear plot (at least in the ideal case, Lineweaver-Burk plot, Fig. 1.5-2a):

$$\frac{1}{v_0} = \frac{K_M}{V_{max} * [S]} + \frac{1}{V_{max}} \qquad [1.5\text{-}21]$$

If $1/v_0$ is plotted vs. 1/[S], then the intersections of this line with abscissa and ordinate allow the determination of K_M and V_{max}.

A disadvantage of the Lineweaver-Burk plot is the accumulation of measuring points near the ordinate (see the markings on the abscissa of Fig. 1.5-2a). Therefore other ways of plotting have been proposed. Hanes used another transformation of the Michaelis-Menten equation:

$$\frac{[S]}{v_0} = \frac{K_M}{V_{max}} + \frac{[S]}{V_{max}} \qquad [1.5\text{-}21a]$$

The plot of $[S]/v_0$ vs. [S] yields a line with the abscissa intersection $-K_M$ and the ordinate intersection K_M/V_{max}. The slope equals $1/V_{max}$ (Fig. 1.5-2 b).

Still another method, the so-called 'direct plot', has been proposed by Eisenthal and Cornish-Bowden. The Michaelis-Menten equation is rearranged as follows:

$$V_{max} = v_0 + \frac{v_0}{[S]} * K_M \qquad [1.5\text{-}21b]$$

For each individual measurement, −[S] is marked on the abscissa and v_0 on the ordinate and a line is drawn through both points. The intersection of these lines has the abscissa value K_M and the ordinate value V_{max} (Fig. 1.5-2 c).

However, the most accurate method is the statistical evaluation of the measurements. In spite of this, the Lineweaver-Burk plot will be used in the following graphical representations, since it is the best known one.

Characterization of enzyme activities: The enzyme activity is defined as the quantity of substrate turned over per time unit in the presence of a given amount of enzyme. Thus the standard dimension would be [mol * sec^{-1}] = katal. For practical reasons, usually the

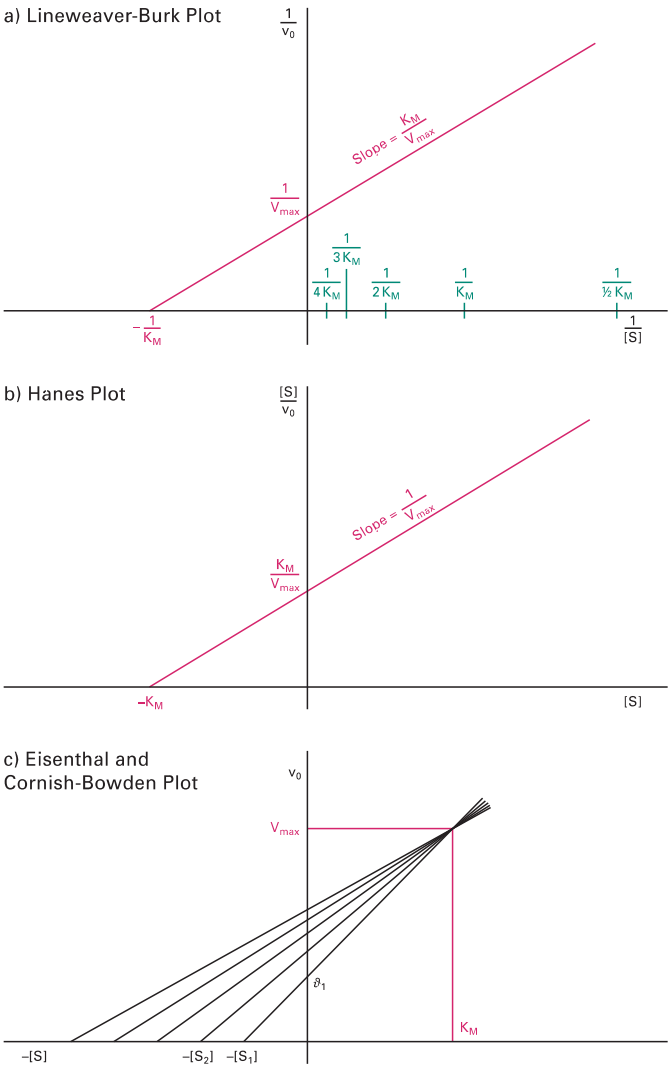


Figure 1.5-2. Linear Plots of an Enzyme Catalyzed Reaction

activity is expressed as [µmol * min^{-1}]. This term is named International Unit (U) if the measurement is performed under standard conditions (with isolated enzymes at conditions that are optimized as much as possible). The specific activity is the enzyme activity per unit of weight, e.g., per mg and is frequently used to characterize the degree of purification of isolated enzymes.

The turnover number of an enzyme is defined as the number of molecules converted by one molecule of enzyme per unit of time if the enzyme is saturated with substrate ($[E_t] = [ES]$). It is identical to the rate constant k_2 and can be calculated from Eq. [1.5-14 b] as $k_2 = V_{max}/[E_t]$. Most turnover numbers are in the range of $1 \dots 10^4$ (see Fig. 10.3-4), the value for catalase is $4 * 10^7$.

Most reactions *in vivo* proceed below the saturation limit of the enzyme, frequently at $[S] = 0.01 \dots 1\ K_M$. By the combination of Eq. [1.5-17], Eq. [1.5-18] and Eq. [1.5-14 a] one obtains

$$v_0 = \frac{k_2}{K_M} * [E] * [S] \qquad [1.5\text{-}22].$$

At low substrate concentration, only a small portion of the enzyme forms an enzyme-substrate complex and $[E] \approx [E_t]$ = constant. The term k_2/K_M indicates how often a contact of enzyme and substrate leads to a reaction and is therefore a measure of the catalytic efficiency. It has an upper limit of ca. 10^9 [l * mol^{-1} * sec^{-1}], when practically every contact leads to a reaction, and the reaction rate is determined by the diffusion speed. The value for catalase ($4 * 10^8$) is one of the highest observed.

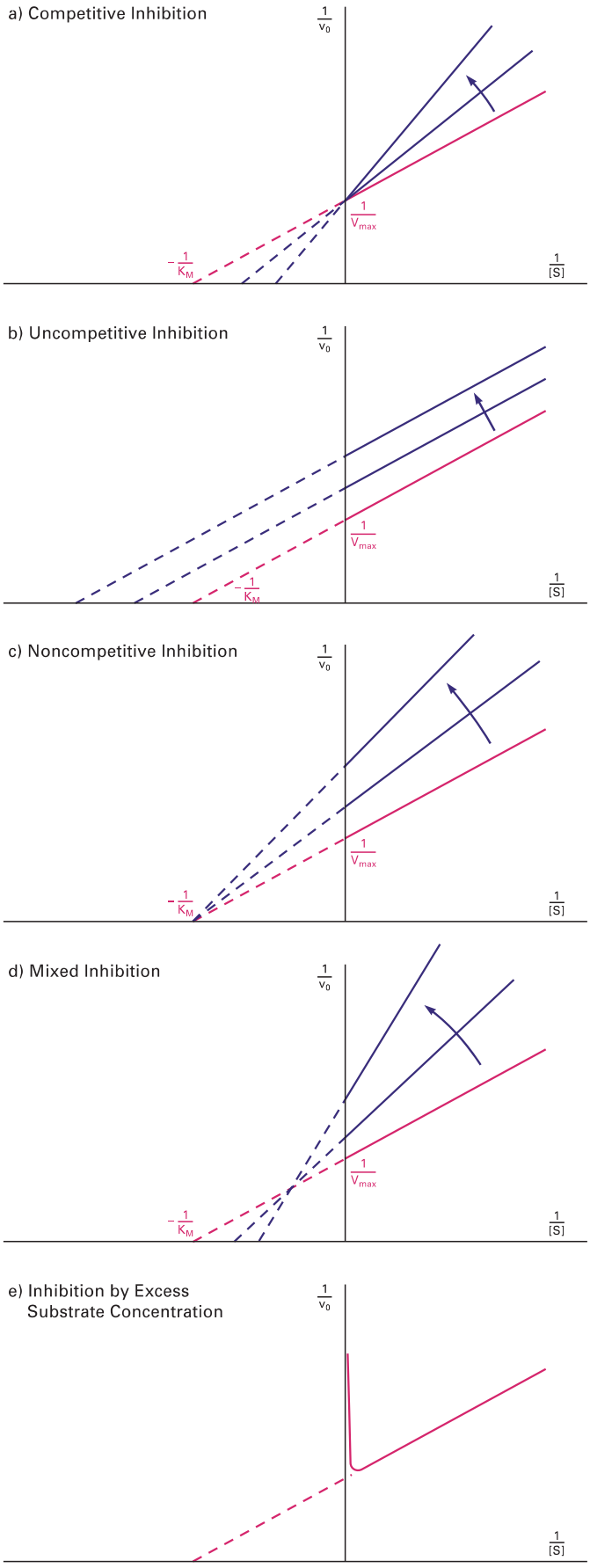


Figure 1.5-3. Lineweaver-Burk Plots of Inhibited Reactions
Red = uninhibited reaction, blue = inhibited reaction, arrow = shift of the plot at increasing inhibitor concentrations.

Inhibition: The mathematical treatment of an inhibited reaction depends on the mechanism of the inhibition. The general principles of inhibition are described in 2.5.2.

Competitive inhibition: The inhibitor competes with the substrate for reversible binding to the active site of the enzyme. The enzyme-substrate and the enzyme-inhibitor complexes are formed with the dissociation constants K_S and K_I, respectively.

$$K_s = \frac{[E] * [S]}{[ES]} \qquad [1.5\text{-}23]$$

$$K_I = \frac{[E] * [I]}{[EI]} \qquad [1.5\text{-}23a]$$

This results in the equation

$$\frac{1}{v_0} = \frac{K_M}{V_{max} * [S]} * \left(1 + \frac{I}{K_I}\right) + \frac{1}{V_{max}} \qquad [1.5\text{-}24]$$

In the Lineweaver-Burk plot, lines obtained at different inhibitor concentration intersect at the ordinate (Fig. 1.5-3a).

Uncompetitive inhibition: The inhibitor reacts reversibly only with the enzyme substrate-complex, but does not affect its formation. The dissociation constant is K'_I.

$$K'_I = \frac{[ES] * [I]}{[ESI]} \qquad [1.5\text{-}25]$$

This yields the equation

$$\frac{1}{v_0} = \frac{K_M}{V_{max} * [S]} + \frac{1}{V_{max}} \left(1 + \frac{[I]}{K'_I}\right) \qquad [1.5\text{-}26]$$

In the Lineweaver-Burk plot, parallel lines are obtained at different inhibitor concentrations (Fig. 1.5-3b).

Noncompetitive and mixed inhibition: If the inhibitor binds both to the enzyme and to the enzyme-substrate complex according to Eqs. [1.5-23a] and [1.5-25] and prevents formation of the product, the following equation results

$$\frac{1}{v_0} = \frac{K_M}{V_{max} * [S]} * \left(1 + \frac{[I]}{K_I}\right) + \frac{1}{V_{max}} * \left(1 + \frac{[I]}{K'_I}\right) \qquad [1.5\text{-}27]$$

If the affinities of the inhibitor to the enzyme and to the enzyme-substrate complex are equal ($K_I = K'_I$), then the lines obtained at different inhibitor concentrations intersect in the Lineweaver-Burk plot at the negative abscissa (K_M remains unchanged, noncompetitive inhibition, Fig. 1.5-3c). Otherwise, they intersect in the second quadrant (left of the ordinate, mixed inhibition, Fig. 1.5-3d).

Inhibition by excessive substrate concentrations. If the reaction velocity decreases at very high substrate concentrations, this results in a Lineweaver-Burk curve bent upwards near the ordinate. This situation is mostly observed in *in vitro* experiments (Fig. 1.5-3e).

Two-substrate reactions: The formulas describing the kinetics are derived from the same assumptions as for one-substrate reactions. Their form depends on the reaction sequence. They involve separate Michaelis constants for the turnover of each substrate.

The Cleland nomenclature system uses the following expressions for the number of the substrates and products of the reaction; 1 – Uni, 2 – Bi, 3 – Ter, 4 – Quad. The substrates are named A, B, C…, the products P, Q, R… and the enzyme species (original state, intermediates and final state) E, F, G… If all components have to combine before the reaction takes place, this is called a sequential reaction. This may take place in an ordered way or at random. If, however, one component leaves the enzyme before the other enters, it is a ping-pong reaction. The mechanisms are schematically drawn in Figure 1.5-4.

The formula for an ordered sequential Bi-Bi reaction is

$$\frac{1}{v_0} = \frac{1}{V_{max}} + \frac{(K_M)_A}{V_{max} * [A]} + \frac{(K_M)_B}{V_{max} * [B]} + \frac{(K_M)_{AB}}{V_{max} * [A] * [B]} \qquad [1.5\text{-}28]$$

The general formula for random sequential reactions is very complicated.

A ping-pong Bi-Bi reaction is described by

$$\frac{1}{v_0} = \frac{1}{V_{max}} + \frac{(K_M)_A}{V_{max} * [A]} * \frac{(K_M)_B}{V_{max} * [B]} \qquad [1.5\text{-}29]$$

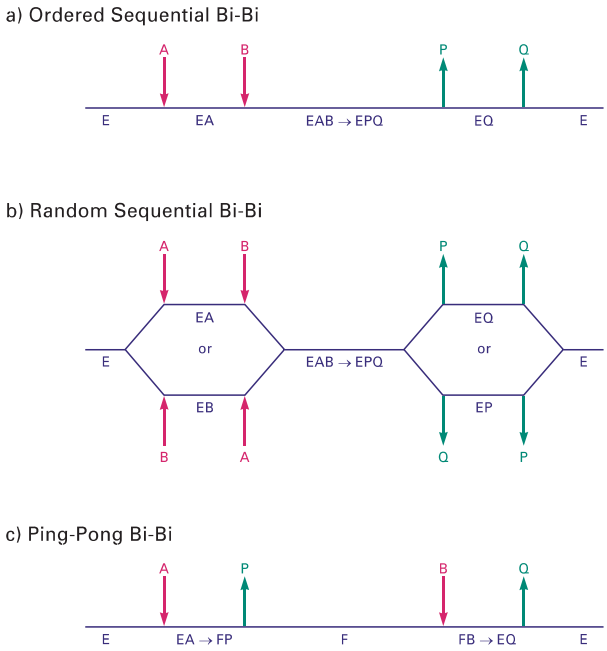


Figure 1.5-4. Types of Two-Substrate-Two-Product (Bi-Bi) Reactions
The enzyme is represented by the horizontal line.

If in second order reactions the concentration of one of the substrates is very much above the respective Michaelis constant, then the terms containing this concentration in Eqs. [1.5-28] and [1.5-29] are practically zero and the equations become identical with Eq. [15.1-21], allowing the same evaluation as with a first order reaction.

If a series of measurements are made in which one substrate is varied while the other is kept constant, then one obtains Lineweaver-Burk plots that formally resemble those obtained with inhibited reactions. However, increasing concentrations of the second substrate shift the lines in the other direction (Fig. 1.5-5). Ordered sequential

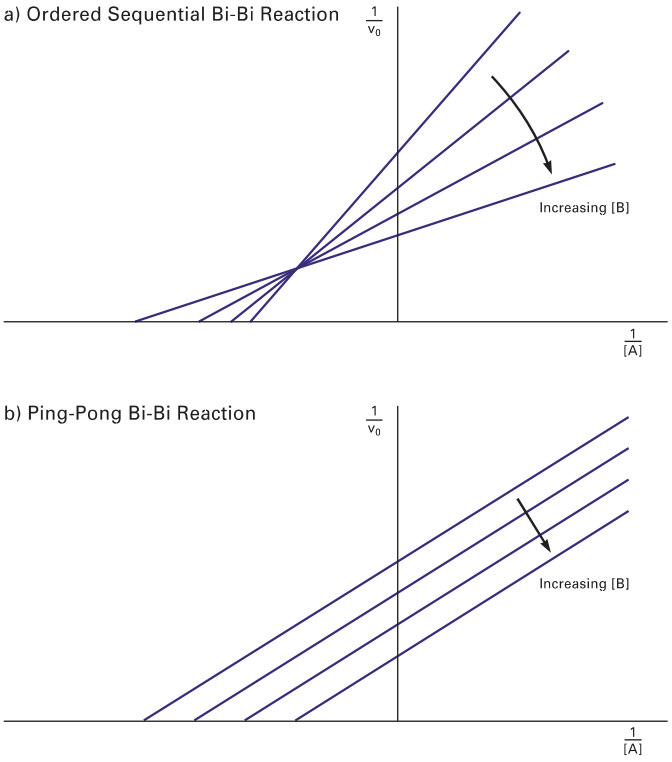


Figure 1.5-5. Lineweaver-Burk Plots of Two-Substrate Reactions
Arrow = shift of the plot when the concentration of the other substrate is raised.

mechanisms yield a series of lines, which intersect left of the ordinate (above or below the abscissa), while ping-pong mechanisms yield parallel lines.

Dependence of reactions on temperature and activation energy: A more refined consideration of the reaction sequence Eq. [1.5-16] shows that only collisions of the reactants above a certain energy level will lead to the formation of complexes, e.g., ES and EP. Also, the reaction ES → EP requires an initial energy input. Thus, the reaction has to cross 'energy hills', which represent metastable states (Fig. 2.4-1). They are called transition complexes $X^{\neq}$ and can either return to the original components or progress towards the products of the reaction, quickly achieving equilibrium in both cases. Among the 'energy hills' mentioned above, the highest one represents obviously the rate determining step of the reaction and has to be the one considered further. (It takes the place of [ES] in the previous equations.) Thus, the equilibrium for formation of this complex can be described analogously to Eq. [1.5-3] by

$$K = \frac{[X^*]}{[A] * [B]} \qquad [1.5\text{-}30]$$

The energy required for its formation is called activation energy $\Delta G^{\neq}$, which can be calculated from this equilibrium by applying Eq. [1.5-4] as

$$\Delta G^{\neq} = -R * T * 2.303 * \log \frac{[X^*]}{[A] * [B]} \qquad [1.5\text{-}31]$$

According to Eq. [1.5-14] the reaction rate for formation of the product(s) from this complex is expressed by $v_0 = k * [X^{\neq}]$. When combining this equation with Eq. [1.5-31], one obtains the following formula for the temperature and $\Delta G^{\neq}$ dependence:

$$v_0 = \text{const.} * [A] * [B] * 10^{-\Delta G^{\neq} / 2{,}303 * R * T} \qquad [1.5\text{-}32]$$

The increase of the reaction rate with rising temperature is limited, however. When the enzyme becomes thermally denatured, the rate drops (Fig. 2.4-4).

Fractal enzyme kinetics: The above considerations assume 'ideal' conditions; purified enzymes, low concentrations, free movement of the reactants. However, *in vivo* the situation is different. Based on a power-law derivation it has been shown that, e.g., restrictions in space require the introduction of non-integer powers > 1 to the concentration terms in Eq. [1.5-17]:

$$\frac{d\,[ES]}{dt} = \alpha_1 * [E]^g * [S]^h - \alpha_{-1} * [ES] - \alpha_2 * [ES] \qquad [1.5\text{-}33]$$

The consecutive equations change analogously. This system is called fractal kinetics. Its main implications are:

- K_M is dependent on the enzyme concentration; it decreases with increasing enzyme concentration.
- The plot of enzyme activity vs. substrate concentration has a tendency towards a sigmoid shape even with monomeric enzymes.
- The velocity of the reaction increases if the movements are, e.g., restricted to surface interfaces (e.g., 3.4.3.2) or to one dimension (e.g., by sliding along nucleic acid strands, 3.8.1.2, 4.2.3.2 or by 'substrate channeling', 3.2.7.1).
- In sequences of reactions, the flux responses are faster and the accumulation of intermediates is lower as compared to the Michealis-Menten assumption.

In some respects, fractal kinetics resemble allosteric situations (2.5.2). Velocity calculations according to this theory have a tendency to yield higher values as according to the Michaelis-Menten theory, which represents a borderline case of a more general treatment, but is still of value for understanding the basic principles of enzyme catalysis.

Literature:

Cornish-Bowden, A., Wharton, C.W. *Enzyme Kinetics*. Oxford: IRL Press; 1988.

Dixon, M., Webb, E.C. *Enzymes.* 3rd Ed. Waltham (MA): Academic Press; 1979.

Fersht, A. *Structure and Mechanism in Protein Science*. New York: W.H. Freeman & Company; 1999.

Freeman (1985). Savageau, M.A. J. Theor. Biol. 1995;176:115–124.

Sigman, D.S., Boyer, P.D. (Eds.). *The Enzymes.* 3rd Ed. Vols. 19 and 20. Waltham (MA): Academic Press; 1990 and 1992.

2 The Cell and Its Contents

Gerhard Michal and Dietmar Schomburg

This chapter presents selected information on the structure and organization of living organisms and their major components to serve as a background for the biochemical text of this book. For more details, refer to biology textbooks.

2.1 Classification of Living Organisms

Life is associated with a number of characteristics such as propagation, metabolism, response to environmental influences, and evolution. Cells are the basic unit of organization for all living beings. Whereas unicellular organisms exist as separate entities, the various cells of multicellular organisms fulfill different functions, and the organism depends on mutual cellular interaction.

There are several systems of classification of living organisms. From a phylogenetic viewpoint, the classification into the three domains; bacteria, archaea and eukarya (which are further subdivided) appears most justified (Table 2.1-1). When common aspects of eubacteria and archaea are discussed, the term prokarya is used.

The metabolic reactions in this book are indicated by colored arrows. Since frequently the occurrence of the reactions is known only for a few species and also in order to prevent an 'overloading' of the figures with too much detail, the arrow colors have been combined into (black) general metabolism, (red) bacteria and archaea, (green) plants, fungi and protists, (blue) animals.

Living organisms exhibit a high degree of order. The sum of all endogenous life processes results in a steady decrease of free energy (1.5.1). Therefore, life can only be kept up by an energy input from the environment, either as light energy or by uptake of oxidizable compounds. Another essential requirement of life is the availability of an adequate carbon source. Living beings can be classified according to the mode of energy uptake and the carbon source used (Table 2.1-2).

During the oxidation of compounds, electrons are released, which have to be taken up by a terminal electron acceptor. Energy wise, oxygen is most favorable (3.11): previous to its appearance in the primeval atmosphere, living organisms had to use other acceptors. This is still the case in oxygen-free habitats (Table 2.1-3).

Literature:

Fox, GE. *et al.*: Science 1980;209:457–463.

Holt, JG. *et al.*: *Bergey's Manual of Descriptive Bacteriology* (9th Ed.). Williams and Wilkins (1994).

Margulis, L., Schwartz, K.V.: *Five Kingdoms*. 2nd. Ed. Freeman (1987).

Woese, CR. *et al.*: Structure of Cells. Proc. Natl. Acad. Sci. USA 1990;87:4576–4579.

2.2 Structure of Cells

2.2.1 Prokaryotic Cells (Fig. 2.2-1)

The genetic information is stored in a single, circular double helix of deoxyribonucleic acid (DNA, 2.6.4). It is located in the central portion of the cell in a densely packed form (nucleoid), but without

Table 2.1-1. Some Typical Properties of Living Organisms (Exceptions exist)

Domains	Bacteria	Archaea	Eukarya			
Kingdoms	Bacteria	Archaea	Protists[1]	Plants	Fungi	Animals
Nucleus	no (common term **prokarya**)		Yes			
Genome	circular, ca.10^6 … 5 * 10^7 kb, extra plasmids		linear, 10^7 … > 10^{11} kb, organized in several chromosomes			
RNA polymerase	one type	several types				
Starting amino acid for translation	formylmethionine	methionine				
Reproduction	binary scission		asexual/sexual			
Cellular organization	unicellular (some are aggregated)		mostly unicellular	multicellular	uni-/multicellular	multicellular
Nutrition (Table 2.1-2)	chemoorganotrophic, photoautotrophic or photoheterotrophic	chemolithotrophic, photoautotrophic or chemoorganotrophic.	chemoorganotrophic or photoautotrophic	photoautotrophic	chemoheterotrophic including saprobiontic	chemoheterotrophic
Size of cells	average 1 … 5 µm, wide variation		average 10 … 100 µm, wide variation			
Cell membranes	rigid, contain peptidoglycans	rigid, without peptidoglycans	rigid or soft	rigid, contain cellulose and lignin	rigid, contain chitin	soft, lipid bilayer only
Internal membranes	no		yes, they enclose organelles /vesicles			

[1] Algae, protozoa, fungi-related. The exact demarcation is under discussion.

Table 2.1-2. Sources of Carbon and of Energy

	Phototrophy (Energy input by light)	Chemotrophy (Energy provided by oxidizable compounds from the environment)
Autotrophy (only CO_2 needed as carbon source)	green plants, some protists, photosynthesizing bacteria (3.12)	Prokarya (mainly archaea). Oxidation of inorganic material (chemolithotrophy, 3.10.7)
Heterotrophy (organic compounds needed as carbon source)	some prokarya	All animals and fungi, non-green plants, many protists and prokarya. Oxidation of organic material (chemoorganothrophy). Included are saprobionts (use decaying organic material) and parasites (feed from living beings)

Table 2.1-3. Terminal Electron Acceptors for Oxidation Reactions

	Atmospheric Oxygen		
	Not required		Required
Energy obtained by	anaerobic respiration (3.10.6)	fermentation (3.10.5)	aerobic respiration (3.11)
Electron acceptors	oxidized external compounds (mostly inorganic)	internally generated compounds	atmospheric O_2
Organisms	anaerobes: part of archaea and bacteria		all other organisms
	facultative anaerobes (bacteria)		

Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology, Second Edition. Edited by Gerhard Michal and Dietmar Schomburg.

a special separation from the rest of the cell. Its replication and the translation of the information into protein structures are described in 3.8 and 4.1.

In prokarya, frequently a number of plasmids may occur, which also consist of circular DNA and replicate independently of the main DNA. They carry only a few genes. Although plasmids are not usually essential for survival, they are involved in DNA transfer during conjugation, provide resistance to antibiotics, etc. Some plasmids can be reversibly integrated into the main DNA (episomes). Similar properties are exhibited by DNA viruses and retroviruses (5.2, 5.4). The translocation of genetic material is not discussed here.

The cytoplasm is a semifluid, concentrated solution of proteins, metabolites, nucleotides, salts, etc. It also contains several thousand ribosome particles involved in translation (4.1.2). It is the site of most metabolic reactions and exchanges material in a controlled way with the environment (3.10.1 … 4).

Prokaryotic cells are surrounded by an envelope (3.10.1, Fig. 2.2-2). It not only has an enclosing and protective function. Rather, a number of metabolic reactions take place at transmembrane proteins (e.g., respiration and ATP synthesis) or at membrane associated proteins. In bacteria, the sequence of membrane components from the interior outwards is:

- The plasma membrane, a lipid bilayer with embedded proteins
- The rigid cell wall, which in the case of bacteria consists of either multiple layers (Gram positive bacteria) or a single layer (Gram negative bacteria) of peptidoglycans (murein).
- An additional outer membrane (only in Gram negative bacteria).
- An additional gelatinous capsule is frequently superimposed on the cell wall. It consists mainly of polysaccharides (polymerized glucose, rhamnose, uronic acids etc.). There may also be mucus layers.

Extensions of the cell envelope are pili and flagella, which provide for cellular contact, conjugation, propulsion, etc.

The composition of an *E. coli* cell by weight is H_2O about 70 %, protein 15 %, DNA 1 %, RNA 6 %, polysaccharides 3 %, lipids 2 % (both are mainly present in the envelope), small organic molecules 1 %, inorganic molecules 1 %.

Mycoplasms are a group of bacteria which lack a cell wall. Among them are the smallest self-reproducing organisms (0.10 … 0.25 µm diameter).

Archaea differ from bacteria by

- A different composition and arrangement of rRNAs
- Differences in the RNA polymerase and in the translation mechanism (Table 2.1-1)
- Different composition of the cellular envelope. e.g., murein (3.10.1) is absent, acylglycerols are replaced by branched chain glycerol ethers (3.4.3.3)
- Unusual pathways of metabolism and habitats (methanogens, 3.10.6.2, halobacteria, 3.12.1, thermophiles, etc.)

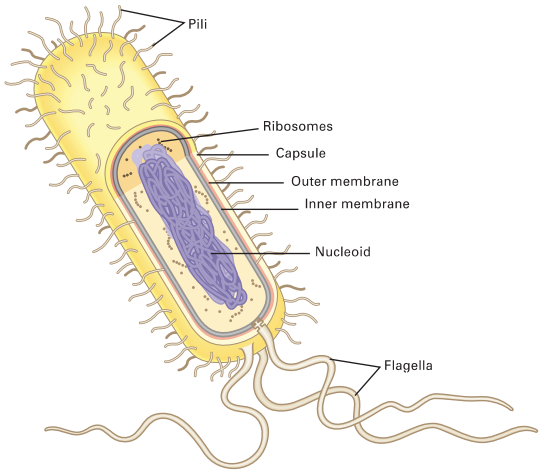


Figure 2.2-1: General Structure of a Bacterial Cell
The colors are for easy differentiation only. After Campbell, N.A.: Biology 4th Ed. Benjamin/Cummings 1996.

2.2.2 General Characteristics of Eukaryotic Cells (Fig. 2.2-3)

Compared to prokaryotic cells, eukaryotic cells exhibit a much more complicated structure. Inside the plasma membrane there are the nucleus and the cytoplasm, which encompass the fluid cytosol and many organelles. These are compartments enclosed by individual membranes, which are devoted to specific functions.

Nucleus: All eukaryotic cells show the presence of a separate nucleus, which contains the major portion of the genetic material of the cell. (The rest is present in mitochondria and chloroplasts, see below.) The nuclear DNA is organized in a number of chromosomes. Each double helix of chromosomal DNA (2.6.3) can be present once (in haploid organisms) or twice (in diploid organisms). During cell division (4.3), the condensed chromosomes arrange themselves separately. Otherwise they are combined with proteins as a ball of chromatin with an elaborate fine structure (2.6.4).

The number of chromosomes present in the various species differs widely (from 4 to >500; humans have 46 in the diploid set). While bacterial genomes contain $< 10^6 \ldots 5 * 10^7$ bp, the diploid set of eukaryotic DNA varies between ca. 10^7 bp (some fungi) and $> 10^{11}$ bp (lungfish, some algae). The diploid human genome contains ca. $5.8 * 10^9$ bp ($2 * 2{,}900{,}000$ bp).

In addition to the DNA the nucleus also contains the nuclear matrix, which is composed of the enzymes and factors required for DNA replication, DNA repair, transcription, and processing of the transcription products (3.9, Chapter 4).

The nucleus is surrounded by a double membrane of lipid bilayers with integrated proteins. Nuclear pores (4.5.3, ca. 125 nm diameter) span the nuclear membrane and enable the transport of proteins, rRNA etc. The inner surface of the nuclear membrane is covered by nuclear lamina, a net of protein fibers which stabilizes the structure

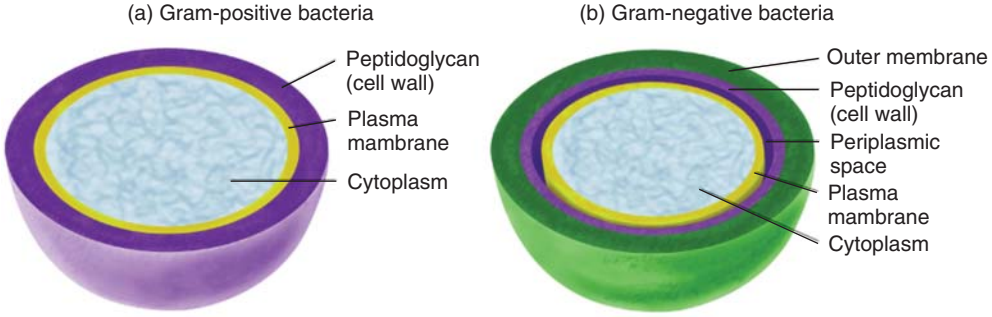


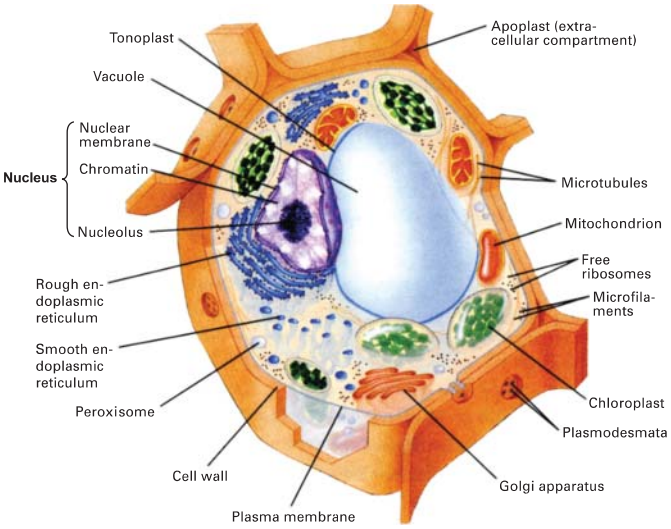
Figure 2.2-2. Structure of the Bacterial Envelope
After Voet, D. and J.G, Pratt, C.W.: Fundamentals of Biochemistry. John Wiley & Sons 2002, Figure 8-14.

and provides attachment points for the chromatin (2.6.4). During cell division, the nuclear membrane dissolves.

Cytosol: Although in eukarya many cytosolic functions have been taken over by specific organelles (see below), a large number of metabolic reactions still take place here. This includes glycolysis (3.1.1) and the synthesis of cytosolic proteins (4.2) and fatty acids (3.4.1.1).

Endoplasmic reticulum (ER, Fig. 2.2-4, Section 4.4.1): This is a labyrinth of membranes, which frequently encompasses half of the total amount of membranes. It consists of a system of sac- and tube-like structures, which locally expand into cisterns. Its internal lumen is connected with the intermembrane space of the nuclear membrane.

Part of the ER is studded on the outside with ribosomes (rough ER), which take part in protein synthesis (4.2). The proteins thus formed enter the ER lumen and are processed mainly by glycosylation (4.4.2). They leave the ER via vesicles (4.5.2). The rough ER is also the site of membrane-phospholipid synthesis, thus providing membrane material for the departing vesicles (3.4.3, 4.4.1).



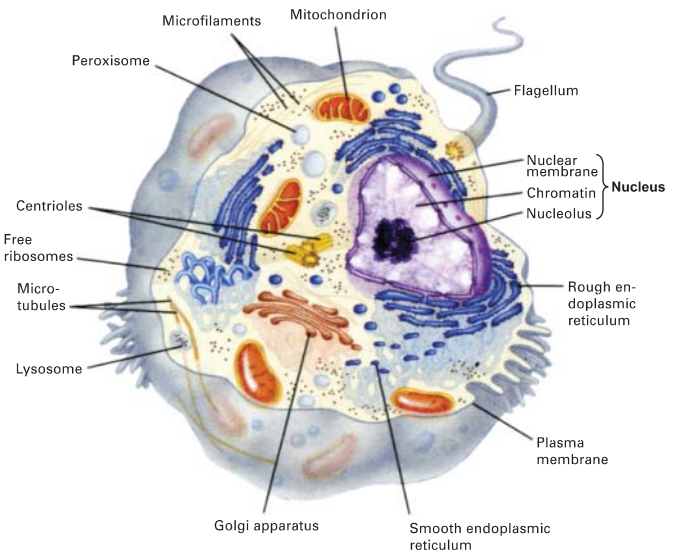


Figure 2.2-3. General Structure of a Plant Cell (Top) and an Animal Cell (Below)
The colors are for easy differentiation only. After Campbell, N.A.: Biology 4th Ed. Benjamin/Cummings 1996.

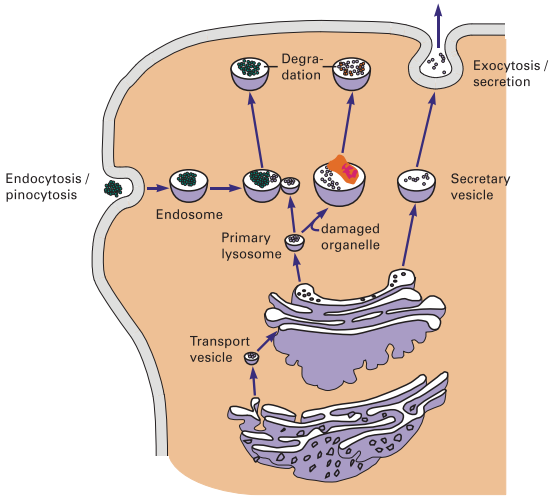


Figure 2.2-4. Interrelationship of Rough Endoplasmic Reticulum, Golgi Apparatus, Lysosomes, Endosomes and Transport Vesicles

The other part of the ER is free of ribosomes (smooth ER). Enzymes of the smooth ER are involved in the synthesis of fatty acids (desaturation, 3.4.1.3), phospholipids (3.4.3), steroids, especially steroid hormones (3.5.4), and other lipids. They are located on the outside of this organelle. The smooth ER plays a role in detoxification by hydroxylation reactions. Its equivalent in muscles (sarcoplasmic reticulum) stores Ca^{2+} ions and is thus involved in the many reactions regulated by this ion (7.4.4). In liver, the smooth ER is also the storage site for glycogen (3.1.2.2).

Golgi apparatus (4.5.2): This organelle consists of stacks of flattened membrane sacs, which are especially numerous in plants (dictyosomes). The properties on both sides of an individual sac and the sacs of the whole stack differ (details in 4.4.2). Their main function is the further processing and sorting of proteins and their export to the final targets. In most cases, these are secreted or membrane proteins. In addition, the Golgi apparatus also produces polysaccharides (e.g., hyaluronic acid, 2.9.2), glycosphingolipids (3.4.4.1) etc.

Lysosomes: Lysosomes are vesicles of 0.1 … 0.8 µm diameter, which are enclosed by a lipid bilayer. They are formed by budding from the *trans* side of the Golgi apparatus (4.4.2). These organelles are filled with many enzymes for polysaccharide, lipid, protein and nucleic acid degradation. They fuse with endosomes containing internalized lipoproteins (6.2.4) or phagocytosed nutrients (6.1.5) and hydrolyze these compounds (Fig. 2.2-4). They act also on intracellular material to be removed and even contribute to the apoptosis (programmed cell death, 7.6, 8.2.5) of their own cell. Lysosomes of special cells (e.g., macrophages) destroy, in the same way, bacteria or viruses as a defense mechanism (8.2.8).

The degradative enzymes of lysosomes exhibit an activity optimum at pH = 5, identical to the pH of the lysosome lumen. It is kept constant by continuous proton pumping into the lysosome lumen. The pH difference to the usual cytosolic pH of 7.0 is a safety measure, since after accidental leakage of some lysosomes, the released enzymes are almost inactive at cytosolic pH. Only a cumulated release from many lysosomes is deleterious to the cell.

Insufficient activity of lysosomal enzymes is the reason for many diseases. Examples are mucopolysaccharidoses (2.9.2), gangliosidoses (3.4.4) and glycogen storage diseases (e.g., Pompe's disease, 3.1.2.5).

Peroxisomes: This is another example of how enclosure into an organelle allows reactions to take place which would otherwise be deleterious to the rest of the cell. Peroxisomes (ca. 0.5 µm diameter) are surrounded by a single membrane. They are generated from components of the cytosol and do not bud from other membranes. The main task of these organelles is the performance of monooxygenase (hydroxylase) or oxidase reactions, which produce hydrogen peroxide (H_2O_2).

This dangerous compound is immediately destroyed by catalase within this enclosed space:

$X\text{-}H + O_2 + H_2O = X\text{-}OH + H_2O_2$ or

$X\text{-}CH_2OH + O_2 = X\text{-}CHO + H_2O_2$, followed by

$2\ H_2O_2 = 2\ H_2O + O_2$.

Reactions of this kind are the oxidative degradation of fatty acids (3.4.1.5), of alcohols etc. In glyoxisomes of plants, the reactions of the glyoxylate cycle (3.1.9.1) take place.

Mitochondria (Fig. 2.2-5): In a typical eukaryotic cell, there are in the order of 2000 of these organelles, which are often of ellipsoidal shape (length ca. 1 … 10 µm). They have a smooth outer membrane and a highly folded inner membrane with numerous invaginations (cristae), which contain most of the membrane-bound enzymes of mitochondrial metabolism. The internal area contains the mitochondrial matrix, while the intermembrane space is the narrow area between both membranes. Since protons are permanently pumped out through the inner membrane, the matrix is more alkaline than the intermembrane space and the cytoplasm (pH 8 vs. ≈ pH 7).

Mitochondria are the site of respiration and ATP synthesis (3.11.4.5), but also of many other central reactions of metabolism, e.g., citrate cycle (3.1.8), fatty acid oxidation (3.4.1.5), glutamine formation (3.2.2.1), and part of the pathway leading to steroid hormones (3.5.5.1, 3.5.8.1). The latter sequence, as well as the initiation of gluconeogenesis (3.1.3.5), are examples of how the site of reactions frequently changes from one organelle to another within a single pathway.

Besides chloroplasts (2.2.3), mitochondria are the only organelles which are equipped with their own (circular) DNA, RNA and ribosomes and thus can perform their own protein synthesis. The components and the mechanism resemble more the bacterial than the eukaryotic system. However, less than 10 % of the mitochondrial proteins are generated by this means, the rest is encoded by nuclear DNA and imported (4.5). Mitochondria reproduce by binary scission similar to bacteria. Their membranes do not exchange material with the rest of the cellular membrane system; the membrane proteins are produced on internal or cytosolic ribosomes. The membranes contain many transport systems resembling those of bacteria. These and other arguments are the base of the endosymbiont theory, which assumes that mitochondria have originated from ingested bacteria (e.g., aerobically living heterotrophs). They provide effective energy production, but are dependent on the host in many aspects.

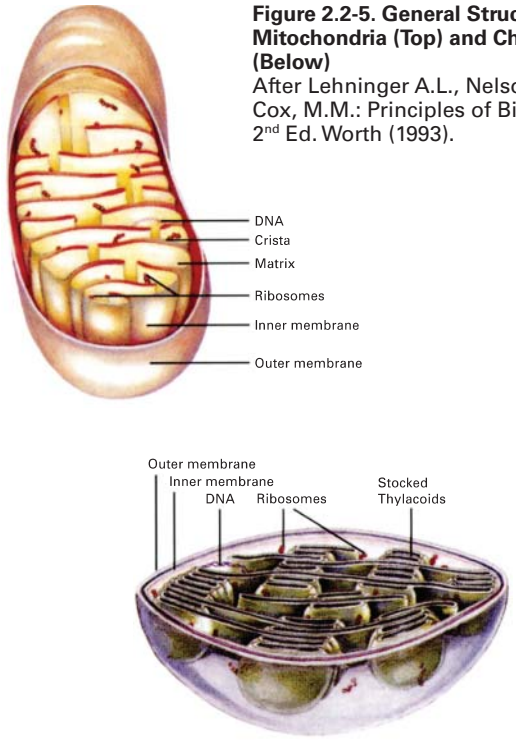


Figure 2.2-5. General Structure of Mitochondria (Top) and Chloroplasts (Below)
After Lehninger A.L., Nelson, D.L., Cox, M.M.: Principles of Biochemistry 2nd Ed. Worth (1993).

Cytoskeleton: The internal cytoskeleton is a general component of eukaryotic cells, but is of special importance in animal cells. Here it is the primary factor for maintenance of the external structure, which is achieved in plant cells by means of the rigid cell wall (2.2.3). The cytoskeleton also provides anchoring points for the organelles and for some enzymes and even contributes to movements either by its own formation and degradation or by motor proteins moving along the filaments.

The cytoskeleton is composed of microtubules, actin filaments and intermediary filaments (Table 2.2-1). More details on their composition, formation and degradation are presented in 6.1.6.

Table 2.2-1. Components of the Cytoskeleton

	Microtubules	Microfilaments	Intermediary Filaments
Components	α/β-tubulin	actin	keratin, vimetin, desmin etc.
Shape	hollow tube of 13 tubulin sequences	2 strings of actin monomers, coiled	variable, frequently coiled coil of α helices
Diameter	25 nm	7 nm	8 … 12 nm
Examples of function	chromosome separation (4.3.5), axonal transport (7.2.6), other movements	muscle contraction (7.4.5), movement of the cytoplasm, cell division (furrow ingression, 4.3.5)	structure stabilization, anchoring of organelles

The hollow microtubules (Fig. 6.1-3) frequently originate from the centrosome, which is located close to the nucleus (4.3.5). The moving parts of cilia and flagella are circular arrangements of tubulin. Attached dynein molecules cause movements of the tubules relative to each other and thus curvature of these cellular annexes (6.1.6).

Actin filaments associate with myosin. The heads of the myosin molecules can 'walk along' the string of actin monomers. The best known action of actin filaments is the muscle contraction (7.4.5), but they are also involved in many other, mostly movement, functions. E.g., the cytoplasmic movement in some algae is caused by the transport of organelles along actin filaments by myosin.

Intermediary filaments are composed of variable units, depending on the particular cell type. In general they have a structural function. They are more long-lasting than the other components of the cytoskeleton, but are also subject to rearrangements. Vimetin occurs in, e.g., endothelial cells and adipocytes and anchors the nucleus and the fat droplets. Desmin filaments keep the Z disks of muscle cells in place (7.4.5). Neurofilaments reinforce the long axons of neuronal cells. The nuclear lamina (see above) consists likewise of intermediary filaments.

2.2.3 Special Structures of Plant Cells (Fig. 2.2-3)

Plant cell walls: Cell walls are an essential factor distinguishing plant and animal cells. They provide stability and prevent an expansion of the cell beyond its fixed size. These secondary walls consist mainly of cellulose (3.1.2.2), hemicelluloses (3.1.6.3) and lignin (3.13.1.2). During an early phase of formation, the primary walls contain considerable amounts of pectate (3.1.5.6) and glycoproteins (4.4.1). On the inside, the cell contents are enclosed by the usual lipid bilayer membrane. Plasmodesmata interconnect neighboring cells and enable the transport of water and metabolites.

Chloroplasts (Fig. 2.2-5): These are lens-shaped organelles of about 1 … 5 µm diameter, which occur only in photosynthesizing green plants. They are enclosed by a double membrane with a thin intermembrane space in between. The interior contains the fluid stroma and a third membrane system surrounding the thylakoid space. This system has the shape of interconnected flat disks, which, in most cases, are stacked on top of each other (grana). The thylakoid membranes are the site of photosynthesis (3.12.1). In photosynthesizing bacteria, their role is taken over by the cytoplasmic membrane.

Due to the permanent pumping of protons into the thylakoid space, its interior is much more acidic than the stroma (pH = 5 vs. pH = 8).