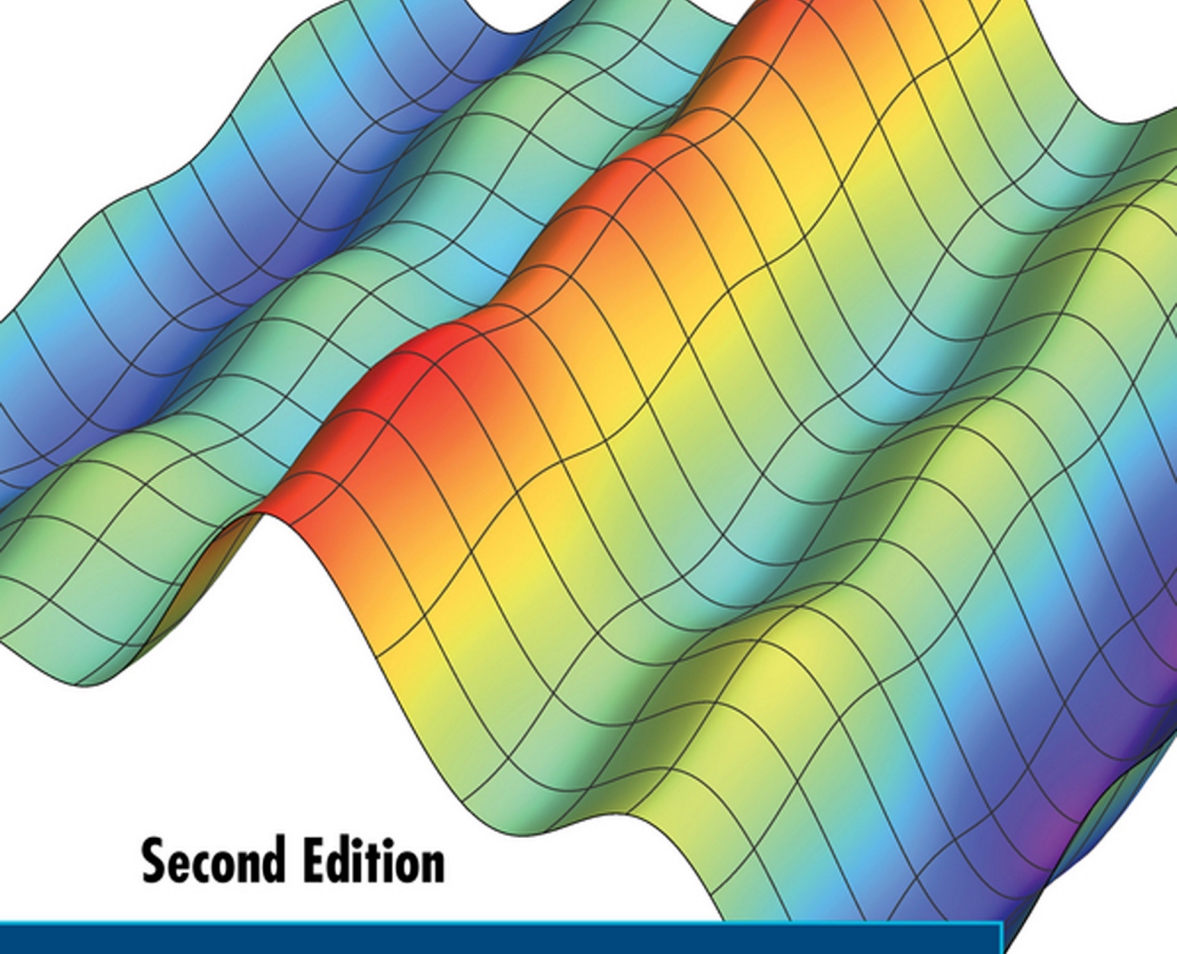




Second Edition

PHYSICAL CHEMISTRY

for the

BIOLOGICAL SCIENCES

Gordon G. Hammes • Sharon Hammes-Schiffer

WILEY

PHYSICAL CHEMISTRY FOR THE BIOLOGICAL SCIENCES

METHODS OF
BIOCHEMICAL ANALYSIS

Volume 55

A complete list of the titles in this series appears at the end of this volume.

PHYSICAL CHEMISTRY FOR THE BIOLOGICAL SCIENCES

SECOND EDITION

Gordon G. Hammes
Sharon Hammes-Schiffer

Part of Wiley Series in Methods of Biochemical Analysis

WILEY

Copyright © 2015 by John Wiley & Sons, Inc. All rights reserved

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

Published simultaneously in Canada

No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning, or otherwise, except as permitted under Section 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the Publisher, or authorization through payment of the appropriate per-copy fee to the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, (978) 750-8400, fax (978) 750-4470, or on the web at www.copyright.com. Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, (201) 748-6011, fax (201) 748-6008, or online at <http://www.wiley.com/go/permissions>.

Limit of Liability/Disclaimer of Warranty: While the publisher and author have used their best efforts in preparing this book, they make no representations or warranties with respect to the accuracy or completeness of the contents of this book and specifically disclaim any implied warranties of merchantability or fitness for a particular purpose. No warranty may be created or extended by sales representatives or written sales materials. The advice and strategies contained herein may not be suitable for your situation. You should consult with a professional where appropriate. Neither the publisher nor author shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages.

For general information on our other products and services or for technical support, please contact our Customer Care Department within the United States at (800) 762-2974, outside the United States at (317) 572-3993 or fax (317) 572-4002.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic formats. For more information about Wiley products, visit our web site at www.wiley.com.

Library of Congress Cataloging-in-Publication Data:

Hammes, Gordon G., 1934–

Physical chemistry for the biological sciences. – Second edition / Gordon G. Hammes,
Sharon Hammes-Schiffer.

pages cm. – (Wiley series in methods of biochemical analysis)

Includes index.

ISBN 978-1-118-85900-1 (cloth)

1. Physical biochemistry. 2. Thermodynamics. 3. Chemical kinetics. 4. Biomolecules–Spectra.
5. Spectrum analysis. I. Hammes-Schiffer, Sharon. II. Title.

QP517.P49H348 2015

612'.01583–dc23

2014043242

Cover image courtesy of G. G. Hammes, S. J. Benkovic, and S. Hammes-Schiffer, *Biochemistry* **50**, 10422 (2011). © 2011 by American Chemical Society.

Typeset in 10/12pt Times-Roman by Laserwords Private Limited, Chennai, India.

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

2 2015

CONTENTS

Preface to First Edition	xv
Preface to Second Edition	xvii
 THERMODYNAMICS	 1
1. Heat, Work, and Energy	3
1.1 Introduction	3
1.2 Temperature	4
1.3 Heat	5
1.4 Work	6
1.5 Definition of Energy	9
1.6 Enthalpy	11
1.7 Standard States	12
1.8 Calorimetry	13
1.9 Reaction Enthalpies	16
1.10 Temperature Dependence of the Reaction Enthalpy	18
References	19
Problems	20
 2. Entropy and Gibbs Energy	 23
2.1 Introduction	23
2.2 Statement of the Second Law	24
2.3 Calculation of the Entropy	26
2.4 Third Law of Thermodynamics	28
2.5 Molecular Interpretation of Entropy	29

2.6	Gibbs Energy	30
2.7	Chemical Equilibria	32
2.8	Pressure and Temperature Dependence of the Gibbs Energy	35
2.9	Phase Changes	36
2.10	Additions to the Gibbs Energy	39
	Problems	40
3.	Applications of Thermodynamics to Biological Systems	43
3.1	Biochemical Reactions	43
3.2	Metabolic Cycles	45
3.3	Direct Synthesis of ATP	49
3.4	Establishment of Membrane Ion Gradients by Chemical Reactions	51
3.5	Protein Structure	52
3.6	Protein Folding	60
3.7	Nucleic Acid Structures	63
3.8	DNA Melting	67
3.9	RNA	71
	References	72
	Problems	73
4.	Thermodynamics Revisited	77
4.1	Introduction	77
4.2	Mathematical Tools	77
4.3	Maxwell Relations	78
4.4	Chemical Potential	80
4.5	Partial Molar Quantities	83
4.6	Osmotic Pressure	85
4.7	Chemical Equilibria	87
4.8	Ionic Solutions	89
	References	93
	Problems	93

CHEMICAL KINETICS 95**5. Principles of Chemical Kinetics 97**

5.1	Introduction	97
5.2	Reaction Rates	99
5.3	Determination of Rate Laws	101
5.4	Radioactive Decay	104
5.5	Reaction Mechanisms	105
5.6	Temperature Dependence of Rate Constants	108
5.7	Relationship Between Thermodynamics and Kinetics	112
5.8	Reaction Rates Near Equilibrium	114
5.9	Single Molecule Kinetics	116
	References	118
	Problems	118

6. Applications of Kinetics to Biological Systems 121

6.1	Introduction	121
6.2	Enzyme Catalysis: The Michaelis–Menten Mechanism	121
6.3	α -Chymotrypsin	126
6.4	Protein Tyrosine Phosphatase	133
6.5	Ribozymes	137
6.6	DNA Melting and Renaturation	142
	References	148
	Problems	149

QUANTUM MECHANICS 153**7. Fundamentals of Quantum Mechanics 155**

7.1	Introduction	155
7.2	Schrödinger Equation	158
7.3	Particle in a Box	159
7.4	Vibrational Motions	162

7.5	Tunneling	165
7.6	Rotational Motions	167
7.7	Basics of Spectroscopy	169
	References	173
	Problems	174
8.	Electronic Structure of Atoms and Molecules	177
8.1	Introduction	177
8.2	Hydrogenic Atoms	177
8.3	Many-Electron Atoms	181
8.4	Born–Oppenheimer Approximation	184
8.5	Molecular Orbital Theory	186
8.6	Hartree–Fock Theory and Beyond	190
8.7	Density Functional Theory	193
8.8	Quantum Chemistry of Biological Systems	194
	References	200
	Problems	201
	SPECTROSCOPY	203
9.	X-ray Crystallography	205
9.1	Introduction	205
9.2	Scattering of X-Rays by a Crystal	206
9.3	Structure Determination	208
9.4	Neutron Diffraction	212
9.5	Nucleic Acid Structure	213
9.6	Protein Structure	216
9.7	Enzyme Catalysis	219
	References	222
	Problems	223

10. Electronic Spectra	225
10.1 Introduction	225
10.2 Absorption Spectra	226
10.3 Ultraviolet Spectra of Proteins	228
10.4 Nucleic Acid Spectra	230
10.5 Prosthetic Groups	231
10.6 Difference Spectroscopy	233
10.7 X-Ray Absorption Spectroscopy	236
10.8 Fluorescence and Phosphorescence	236
10.9 RecBCD: Helicase Activity Monitored by Fluorescence	240
10.10 Fluorescence Energy Transfer: A Molecular Ruler	241
10.11 Application of Energy Transfer to Biological Systems	243
10.12 Dihydrofolate Reductase	245
References	247
Problems	248
11. Circular Dichroism, Optical Rotary Dispersion, and Fluorescence Polarization	253
11.1 Introduction	253
11.2 Optical Rotary Dispersion	254
11.3 Circular Dichroism	256
11.4 Optical Rotary Dispersion and Circular Dichroism of Proteins	257
11.5 Optical Rotation and Circular Dichroism of Nucleic Acids	259
11.6 Small Molecule Binding to DNA	260
11.7 Protein Folding	263
11.8 Interaction of DNA with Zinc Finger Proteins	266
11.9 Fluorescence Polarization	267
11.10 Integration of HIV Genome Into Host Genome	269
11.11 α -Ketoglutarate Dehydrogenase	270
References	272
Problems	273

12. Vibrations in Macromolecules	277
12.1 Introduction	277
12.2 Infrared Spectroscopy	278
12.3 Raman Spectroscopy	279
12.4 Structure Determination with Vibrational Spectroscopy	281
12.5 Resonance Raman Spectroscopy	283
12.6 Structure of Enzyme–Substrate Complexes	286
12.7 Conclusion	287
References	287
Problems	288
13. Principles of Nuclear Magnetic Resonance and Electron Spin Resonance	289
13.1 Introduction	289
13.2 NMR Spectrometers	292
13.3 Chemical Shifts	293
13.4 Spin–Spin Splitting	296
13.5 Relaxation Times	298
13.6 Multidimensional NMR	300
13.7 Magnetic Resonance Imaging	306
13.8 Electron Spin Resonance	306
References	310
Problems	310
14. Applications of Magnetic Resonance to Biology	315
14.1 Introduction	315
14.2 Regulation of DNA Transcription	315
14.3 Protein–DNA Interactions	318
14.4 Dynamics of Protein Folding	320
14.5 RNA Folding	322
14.6 Lactose Permease	325

14.7	Proteasome Structure and Function	328
14.8	Conclusion	329
	References	329
STATISTICAL MECHANICS		331
15.	Fundamentals of Statistical Mechanics	333
15.1	Introduction	333
15.2	Kinetic Model of Gases	333
15.3	Boltzmann Distribution	338
15.4	Molecular Partition Function	343
15.5	Ensembles	346
15.6	Statistical Entropy	349
15.7	Helix-Coil Transition	350
	References	353
	Problems	354
16.	Molecular Simulations	357
16.1	Introduction	357
16.2	Potential Energy Surfaces	358
16.3	Molecular Mechanics and Docking	364
16.4	Large-Scale Simulations	365
16.5	Molecular Dynamics	367
16.6	Monte Carlo	373
16.7	Hybrid Quantum/Classical Methods	373
16.8	Helmholtz and Gibbs Energy Calculations	375
16.9	Simulations of Enzyme Reactions	376
	References	379
	Problems	379

SPECIAL TOPICS **383**

17. Ligand Binding to Macromolecules **385**

17.1	Introduction	385
17.2	Binding of Small Molecules to Multiple Identical Binding Sites	385
17.3	Macroscopic and Microscopic Equilibrium Constants	387
17.4	Statistical Effects in Ligand Binding to Macromolecules	389
17.5	Experimental Determination of Ligand Binding Isotherms	392
17.6	Binding of Cro Repressor Protein to DNA	395
17.7	Cooperativity in Ligand Binding	397
17.8	Models for Cooperativity	402
17.9	Kinetic Studies of Cooperative Binding	406
17.10	Allosterism	408
	References	412
	Problems	412

18. Hydrodynamics of Macromolecules **415**

18.1	Introduction	415
18.2	Frictional Coefficient	415
18.3	Diffusion	418
18.4	Centrifugation	421
18.5	Velocity Sedimentation	422
18.6	Equilibrium Centrifugation	424
18.7	Preparative Centrifugation	425
18.8	Density Centrifugation	427
18.9	Viscosity	428
18.10	Electrophoresis	429
18.11	Peptide-Induced Conformational Change of a Major Histocompatibility Complex Protein	432
18.12	Ultracentrifuge Analysis of Protein–DNA Interactions	434
	References	435
	Problems	435

19. Mass Spectrometry	441
19.1 Introduction	441
19.2 Mass Analysis	441
19.3 Tandem Mass Spectrometry (MS/MS)	445
19.4 Ion Detectors	445
19.5 Ionization of the Sample	446
19.6 Sample Preparation/Analysis	449
19.7 Proteins and Peptides	450
19.8 Protein Folding	452
19.9 Other Biomolecules	455
References	455
Problems	456
 APPENDICES	 457
Appendix 1. Useful Constants and Conversion Factors	459
Appendix 2. Structures of the Common Amino Acids at Neutral pH	461
Appendix 3. Common Nucleic Acid Components	463
Appendix 4. Standard Gibbs Energies and Enthalpies of Formation at 298 K, 1 atm, pH 7, and 0.25 M Ionic Strength	465
Appendix 5. Standard Gibbs Energy and Enthalpy Changes for Biochemical Reactions at 298 K, 1 atm, pH 7.0, pMg 3.0, and 0.25 M Ionic Strength	467
Appendix 6. Introduction to Electrochemistry	469
A6-1 Introduction	469
A6-2 Galvanic Cells	469
A6-3 Standard Electrochemical Potentials	471
A6-4 Concentration Dependence of the Electrochemical Potential	472
A6-5 Biochemical Redox Reactions	473
References	473
Index	475

Biology is the study of living species. The historic origin of biology is descriptive in nature, a classification and description of the various biological species. Modern biology is far different and seeks to understand living phenomena on a molecular basis. The incredible amount of information available and the databases of this information are staggering, the most obvious example being the nucleotide sequence of the human genome. In essence, biology has moved from a qualitative to a quantitative science. Inevitably, this requires a theoretical framework and associated mathematics. Physical chemistry provides this framework for molecular structure and chemical reactions, the components of all biological systems that ultimately must be understood.

Traditionally, physical chemistry has been a major training component for chemists, but not for biologists. This has been attributed to the relatively sophisticated mathematical underpinnings of rigorous physical chemistry. However, the concepts of physical chemistry can be understood and applied to biology with a minimum of mathematics.

This volume attempts to present physical chemistry in conceptual terms using mathematics only at an upper level of elementary calculus, a level required for all science students. Nevertheless, the approach is quantitative in nature, with explicit calculations and numerical problems. Examples from biology are used to illustrate the principles, and problems are appended at the end of each chapter. This book is intended to serve as a one-semester introduction to physical chemistry for undergraduate biology majors and as a refresher course for first-year graduate students. This book combines two volumes published earlier, *Thermodynamics and Kinetics for the Biological Sciences* and *Spectroscopy for the Biological Sciences*. These two books have been integrated with some additions and modification. The most notable addition is a chapter on the hydrodynamics of macromolecules. Hydrodynamics is the basis of several important laboratory techniques used in molecular biology, and understanding the underlying concepts will permit better use of the methods and development of new methods.

We begin with a discussion of thermodynamics, a subject that provides a convenient framework for all equilibrium phenomena. This is followed by chemical kinetics, the quantitative description of the time dependence of chemical reactions. For both subjects, multiple applications to biology are presented. The concepts associated with spectroscopy and structure determination are then considered. These topics deal with the molecular nature of matter and the techniques used to characterize molecules and their interactions. The concluding section of the book includes the important subjects of ligand binding to macromolecules, hydrodynamics, and mass

spectrometry. The coverage of this book represents the minimal knowledge that every biologist should have to understand biological phenomenon in molecular terms (in my opinion!).

I am indebted to my colleagues at Duke for their encouragement and assistance. In particular, Professors Jane and David Richardson, Lorena Beese, Leonard Spicer, Terrance Oas, Michael Fitzgerald, and Harvey Sage who have provided vital expertise. A special thanks also goes to Darla Henderson who as a Wiley editor has provided both encouragement and professional assistance in the preparation of this volume. As always, my wife Judy has provided her much appreciated (and needed) support.

GORDON G. HAMMES
Duke University
Durham, NC, USA

Preface to Second Edition

The impetus for preparing the second edition was twofold. First, the material in the first edition was brought up to date. Although the argument can be made that the principles of physical chemistry are timeless, new applications continually appear. We have tried to ensure that interested students will have access to the most recent developments in the areas covered in this book. Second, with the addition of a co-author, we have significantly expanded and upgraded some of the theoretical aspects of this book. The flavor of the first edition has been retained: students in the biological sciences can still obtain a working knowledge of physical chemistry without utilizing advanced calculus. However, the landscape has changed. Calculus, and even advanced calculus, is now routinely taught in high school so that many more college students have an understanding of advanced calculus. Also research in the biological sciences now includes many more applications of theory relative to ten years ago.

More specifically, five new chapters have been added. The first deals with some of the advanced aspects of thermodynamics and makes use of multivariable calculus. Two of the chapters discuss quantum mechanics in much more detail and at a higher level than the first edition. The additions include a discussion of hydrogen tunneling, as well as a chapter on atomic and molecular electronic structure, with brief treatments of Hartree-Fock and density functional theory. The last two new chapters discuss statistical mechanics. One chapter deals with the fundamentals of the subject, and the other discusses computer simulations, with an extensive treatment of molecular dynamics. Finally, an appendix has been added to introduce the fundamentals of electrochemistry.

As a result of these changes, the second edition contains more material than can be covered in a one semester course. However, the instructor can pick and choose the material to be included for such a course. In fact, this text is suitable for a traditional two semester physical chemistry course. Although a few traditional subjects are not covered, there is more than enough material for two semesters. We have intentionally not designed this text to be encyclopedic in nature to make it more accessible to students for self-study.

We are grateful to a number of people for their assistance in reviewing specific aspects of the book. These people include Professor Nicholas Winograd (Pennsylvania State University), Professor Terrance Oas (Duke University), and Professor Leonard Spicer (Duke University). We again want to thank Professors Jane and David Richardson for the marvelous color plates which have been retained from the first edition. Special thanks are due to Dr. Joshua Layfield, who prepared most of the figures for the new material in the second edition and provided valuable

insights. We also want to acknowledge the support of our spouses, Judy and Peter, who have provided much needed patience and encouragement in this father-daughter endeavor.

GORDON G. HAMMES
Duke University
Durham, NC, USA

SHARON HAMMES-SCHIFFER
University of Illinois
Champaign, IL, USA

THERMODYNAMICS

Heat, Work, and Energy

1.1 INTRODUCTION

Thermodynamics is deceptively simple or exceedingly complex, depending on how you approach it. In this book, we will be concerned with the principles of thermodynamics that are especially useful in thinking about biological phenomena. The emphasis will be on concepts, with a minimum of mathematics. Perhaps an accurate description might be rigor without *rigor mortis*. This may cause some squirming in the graves of thermodynamic purists, but the objective is to provide a foundation for researchers in experimental biology to use thermodynamics. This includes cell biology, microbiology, molecular biology, and pharmacology, among others. A more advanced treatment of some aspects of thermodynamics is presented in Chapter 4. Excellent texts are available that present a more complete exposition of thermodynamics (cf. Refs. 1–3).

In point of fact, thermodynamics can provide a useful way of thinking about biological processes and is indispensable when considering molecular and cellular mechanisms. For example, what reactions and coupled physiological processes are possible? What are the allowed mechanisms involved in cell division or in protein synthesis? What are the thermodynamic considerations that cause proteins, nucleic acids, and membranes to assume their active structures? It is easy to postulate biological mechanisms that are inconsistent with thermodynamic principles—but just as easy to postulate those that are consistent. Consequently, no active researcher in biology should be without a rudimentary knowledge of the principles of thermodynamics. The ultimate goal of this exposition is to understand what determines equilibrium in biological systems and how these equilibrium processes can be coupled together to produce living systems, even though we recognize that living organisms are not at equilibrium. Thermodynamics provides a unifying framework for diverse systems in biology. Both a qualitative and a quantitative understanding are important and will be developed.

The beauty of thermodynamics is that a relatively small number of postulates can be used to develop the entire subject. Perhaps the most important part of this development is to be very precise with regard to concepts and definitions, without

getting bogged down with mathematics. Thermodynamics is a macroscopic theory, not molecular. As far as thermodynamics is concerned, molecules need not exist. However, we will not be purists in this regard: If molecular descriptions are useful for understanding or introducing concepts, they will be used. We will not hesitate to give molecular descriptions of thermodynamic results, but we should recognize that these interpretations are not inherent in thermodynamics itself. It is important to note, nevertheless, that large collections of molecules are assumed so that their behavior is governed by Boltzmann statistics; that is, the normal thermal energy distribution is assumed. This is almost always the case in practice. Furthermore, thermodynamics is concerned with time-independent systems, that is, systems at equilibrium. Thermodynamics has been extended to nonequilibrium systems, but we will not be concerned with the formal development of this subject here.

The first step is to define the *system*. A thermodynamic system is simply that part of the universe in which we are interested. The only caveat is that the system must be large relative to molecular dimensions. The system could be a room, it could be a beaker, it could be a cell, etc. An *open system* can exchange energy and matter across its boundaries, for example, a cell or a room with open doors and windows. A *closed system* can exchange energy but not matter, for example, a closed room or box. An *isolated system* can exchange neither energy nor matter, for example, the universe or, approximately, a closed Dewar flask. We are free to select the system as we choose, but it is very important that we specify what it is. This will be illustrated as we proceed. The *properties* of a system are any measurable quantities characterizing the system. Properties are either *extensive*, proportional to the amount of material in the system, or *intensive*, independent of the amount of material. Examples of extensive properties are mass and volume. Examples of intensive properties are temperature, pressure, and color.

1.2 TEMPERATURE

We are now ready to introduce three important concepts: temperature, heat, and work. None of these are unfamiliar, but we must define them carefully so that they can be used as we develop thermodynamics.

Temperature is an obvious concept, as it simply measures how hot or cold a system is. We will not belabor its definition and will simply assert that thermodynamics requires a unique temperature scale, namely, the Kelvin temperature scale. The Kelvin temperature scale is related to the more conventional Celsius temperature scale by the definition

$$T_{\text{Kelvin}} = T_{\text{Celsius}} + 273.16 \quad (1-1)$$

Although the temperature on the Celsius scale is referred to as “degrees Celsius,” by convention degrees are not stated on the Kelvin scale. For example, a temperature of 100 °C is 373 K. (Thermodynamics is entirely logical—some of the conventions used are not.) The definition of *thermal equilibrium* is very simple: When two systems are at the same temperature, they are at thermal equilibrium.

1.3 HEAT

Heat flows across the system boundary during a change in the state of the system because a temperature difference exists between the system and its surroundings. We know of many examples of heat: Some chemical reactions produce heat, such as the combustion of gas and coal. Reactions in cells can produce heat. By convention, heat flows from higher temperature to lower temperature. This fixes the sign of the heat change. It is important to note that this is a convention and is not required by any principle. For example, if the temperature of the surroundings decreases, heat flows to the system, and the sign of the heat change is positive (+). A simple example will illustrate this sign convention as well as the importance of defining the system under consideration.

Consider two beakers of the same size filled with the same amount of water. In one beaker, A, the temperature is 25 °C, and in the other beaker, B, the temperature is 75 °C. Let us now place the two beakers in thermal contact and allow them to reach thermal equilibrium (50 °C). This situation is illustrated in Figure 1-1. If the system is defined as A, the temperature of the system increases, so the heat change is positive. If the system is defined as B, the temperature of the system decreases, so the heat change is negative. If the system is defined as A and B, no heat flow occurs across the boundary of the system, so the heat change is zero! This illustrates how important it is to define the system before asking questions about what is occurring.

The heat change that occurs is proportional to the temperature difference between the initial and final states of the system. This can be expressed mathematically as

$$q = C(T_f - T_i) \quad (1-2)$$

where q is the heat change, the constant C is the *heat capacity*, T_f is the final temperature, and T_i is the initial temperature. This relationship assumes that the heat capacity is constant, independent of the temperature. In point of fact, the heat capacity often changes as the temperature changes, so that a more precise definition puts this relationship in differential form:

$$dq = C dT \quad (1-3)$$

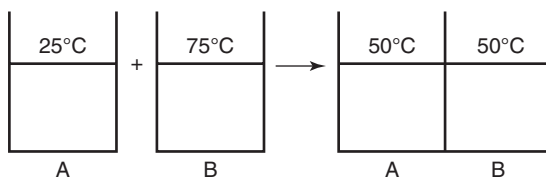


FIGURE 1-1. Illustration of the establishment of thermal equilibrium and importance of defining the *system* carefully. Two identical vessels filled with the same amount of liquid, but at different temperatures, are placed in contact and allowed to reach thermal equilibrium. A discussion of this figure is given in the text.

Note that the heat change and the heat capacity are extensive properties—the larger the system, the larger the heat capacity and the heat change. Temperature, of course, is an intensive property.

1.4 WORK

The definition of *work* is not as simple as that for heat. Many different forms of work exist, for example, mechanical work, such as muscle action, and electrical work, such as ions crossing charged membranes. We will use a rather artificial, but very general, definition of work that is easily understood. Work is a quantity that can be transferred across the system boundary and can always be converted to lifting and lowering a weight in the surroundings. By convention, work done on a system is positive: this corresponds to lowering the weight in the surroundings.

You may recall that mechanical work, w , is defined as the product of the force in the direction of movement, F_x , times the distance moved, x , or in differential form

$$dw = F_x dx \quad (1-4)$$

Therefore, the work to lower a weight is $-mgh$, where m is the mass, g is the gravitational constant, and h is the distance the weight is lowered. This formula is generally useful: for example, mgh is the work required for a person of mass m to walk up a hill of height h . The work required to stretch a muscle could be calculated with Eq. (1-4) if we knew the force required and the distance the muscle was stretched. Electrical work, for example, is equal to $-EIt$, where E is the electromotive force, I is the current, and t is the time. In living systems, membranes often have potentials (voltages) across them. In this case, the work required for an ion to cross the membrane is $-zF\Psi$, where z is the charge of the ion, F is the Faraday (96,489 coulombs/mole), and Ψ is the potential. A specific example is the cotransport of Na^+ and K^+ , Na^+ moving out of the cell and K^+ moving into the cell. A potential of -70 mV is established on the inside so that the electrical work required to move a mole of K^+ ions to the inside is $-(1)(96,489)(0.07) = -6750$ Joules. ($\Psi = \Psi_{\text{outside}} - \Psi_{\text{inside}} = +70$ mV.) The negative sign means that work is done by the system.

Although not very biologically relevant, we will now consider in some detail pressure–volume work, or $P - V$ work. This type of work is conceptually easy to understand, and calculations are relatively easy. The principles discussed are generally applicable to more complex systems, such as those encountered in biology. As a simple example of $P - V$ work, consider a piston filled with a gas, as pictured in Figure 1-2. In this case, the force is equal to the external pressure, P_{ex} , times the area, A , of the piston face, so the infinitesimal work can be written as

$$dw = -P_{\text{ex}} A dx = -P_{\text{ex}} dV \quad (1-5)$$

If the piston is lowered, work is done on the system and is positive, whereas if the piston is raised, work is done by the system and is negative. Note that the work done

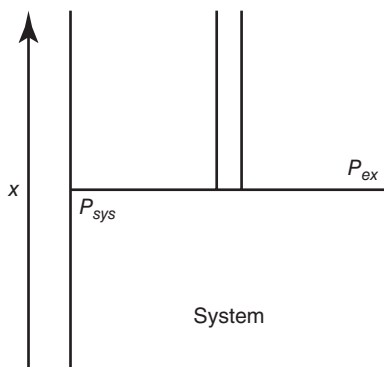


FIGURE 1-2. Schematic representation of a piston pushing on the system. P_{ex} is the external pressure and P_{sys} is the pressure of the system.

on or by the system by lowering or raising the piston depends on what the external pressure is. Therefore, the work can have any value from 0 to ∞ , depending on how the process is done. This is a very important point: the work associated with a given change in state depends on *how* the change in state is carried out.

The idea that work depends on how the process is carried out can be illustrated further by considering the expansion and compression of a gas. The $P - V$ isotherm for an ideal gas is shown in Figure 1-3. An ideal gas is a gas that obeys the ideal gas law, $PV = nRT$ (n is the number of moles of gas and R is the gas constant). The behavior of most gases at moderate pressures is well described by this relationship. Let us consider the expansion of the gas from P_1, V_1 to P_2, V_2 . If this expansion is done with the external pressure equal to zero, that is, into a vacuum, the work is zero. Clearly, this is the minimum amount of work that can be done for this change in state. Let us now carry out the same expansion with the external pressure equal to P_2 . In this case, the work is

$$w = - \int_{V_1}^{V_2} P_{ex} dV = -P_2(V_2 - V_1) \quad (1-6)$$

which is the striped area under the $P - V$ curve. The expansion can be broken into stages; for example, first expand the gas with $P_{ex} = P_3$ followed by $P_{ex} = P_2$, as shown in Figure 1-3. The work done by the system is then the sum of the two rectangular areas under the curve. It is clear that as the number of stages is increased, the magnitude of the work done increases. The maximum work that can be done by the system is when the external pressure is set equal to the pressure of the system minus a small differential pressure, dP , throughout the expansion. This can be expressed as

$$w_{\max} = - \int_{V_1}^{V_2} P dV \quad (1-7)$$

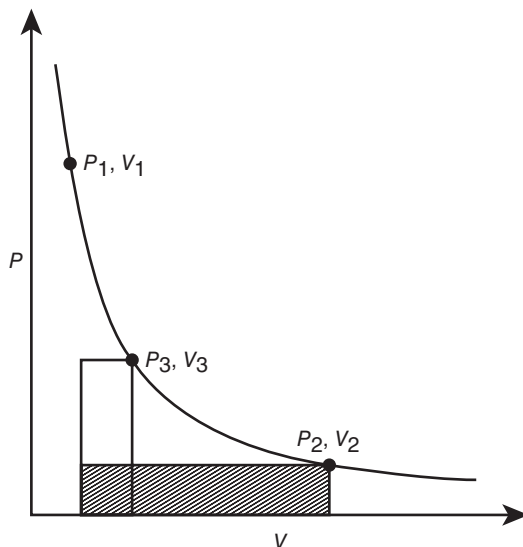


FIGURE 1-3. A P - V isotherm for an ideal gas. The narrow rectangle with both hatched and open areas is the work done in going from P_1, V_1 to P_3, V_3 with an external pressure of P_3 . The hatched area is the work done by the system in going from P_1, V_1 to P_2, V_2 with an external pressure of P_2 . The maximum amount of work done by the system for this change in state is the area under the curve between P_1, V_1 and P_2, V_2 .

By a similar reasoning process, it can be shown that for a compression the minimum work done on the system is

$$w_{\min} = - \int_{V_2}^{V_1} P dV \quad (1-8)$$

This exercise illustrates two important points. First, it clearly shows that the work associated with a change in state depends on how the change in state is carried out. Second, it demonstrates the concept of a *reversible path*. When a change in state is carried out such that the surroundings and the system are not at equilibrium only by an infinitesimal amount, in this case dP , during the change in state, the process is called reversible. The concept of reversibility is only an ideal—it cannot be achieved in practice. Obviously, we cannot really carry out a change in state with only an infinitesimal difference between the pressures of the system and surroundings. We will find this concept very useful, nevertheless.

Now let us think about a cycle whereby an expansion is carried out followed by a compression that returns the system back to its original state. If this is done as a one-stage process in each case, the total work can be written as

$$w_{\text{total}} = w_{\text{exp}} + w_{\text{comp}} \quad (1-9)$$

or

$$w_{\text{total}} = -P_2(V_2 - V_1) - P_1(V_1 - V_2) \quad (1-10)$$

or

$$w_{\text{total}} = (P_1 - P_2)(V_2 - V_1) > 0 \quad (1-11)$$

In this case, net work has been done on the system. For a reversible process, however, the work associated with compression and expansion is

$$w_{\text{exp}} = - \int_{V_1}^{V_2} P dV \quad (1-12)$$

and

$$w_{\text{comp}} = - \int_{V_2}^{V_1} P dV \quad (1-13)$$

so that the total work for the cycle is equal to zero. Indeed, for reversible cycles the net work is always zero.

To summarize this discussion of the concept of work, the work done on or by the system depends on how the change in state of the system occurs. In the real world, changes in state always occur irreversibly, but we will find the concept of a reversible change in state to be very useful.

Heat changes also depend on how the process is carried out. Generally, a subscript is appended to q , for example, q_P and q_V for heat changes at constant pressure and volume, respectively. As a case in point, the heat change at constant pressure is greater than that at constant volume if the temperature of a gas is raised. This is because not only must the temperature be raised, but the gas must also be expanded.

Although this discussion of gases seems far removed from biology, the concepts and conclusions reached are quite general and can be applied to biological systems. The only difference is that exact calculations are usually more difficult. It is useful to consider why this is true. In the case of ideal gases, a simple equation of state is known, $PV = nRT$, that is obeyed quite well by real gases under normal conditions. This equation is valid because gas molecules, on average, are quite far apart and their energetic interactions can be neglected. Collisions between gas molecules can be approximated as billiard balls colliding. This situation obviously does not prevail in liquids and solids where molecules are close together and the energetics of their interactions cannot be neglected. Consequently, simple equations of state do not exist for liquids and solids.

1.5 DEFINITION OF ENERGY

The first law of thermodynamics is basically a definition of the energy change associated with a change in state. It is based on the experimental observation that heat and work can be interconverted. Probably the most elegant demonstration of this is the experimental work of James Prescott Joule in the late 1800s. He carried out experiments in which he measured the work necessary to turn a paddle wheel in water and the concomitant rise in temperature of the water. With this rather primitive experiment, he was able to calculate the conversion factor between work and heat with

amazing accuracy, namely, to within 0.2%. The first law states that the energy change, ΔE , associated with a change in state is

$$\Delta E = q + w \quad (1-14)$$

Furthermore, the energy change is the same regardless of how the change in state is carried out. In this regard, energy clearly has quite different properties than heat and work. This is true for both reversible and irreversible processes. Because of this property, the energy (usually designated the internal energy in physical chemistry textbooks) is called a *state function*. State functions are extremely important in thermodynamics, both conceptually and practically.

Obviously we cannot prove the first law, as it is a basic postulate of thermodynamics. However, we can show that without this law events could occur that are contrary to our experience. Assume, for example, that the energy change in going from state 1 to state 2 is greater than the negative of that for going from state 2 to 1 because the changes in state are carried out differently. We could then cycle between these two states and produce energy as each cycle is completed, essentially making a perpetual motion machine. We know that such machines do not exist, consistent with the first law. Another way of looking at this law is as a statement of the conservation of energy.

It is important that thermodynamic variables are not just hypothetical—we must be able to relate them to laboratory experience, that is, to measure them. Thermodynamics is developed here for practical usage. Therefore, we must be able to relate the concepts to what can be done in the laboratory. How can we measure energy changes? If we only consider $P - V$ work, the first law can be written as

$$\Delta E = q - \int_{V_1}^{V_2} P_{\text{ex}} dV \quad (1-15)$$

If the change in state is measured at constant volume, then

$$\Delta E = q_V \quad (1-16)$$

At first glance, it may seem paradoxical that a state function, the energy change, is equal to a quantity whose magnitude depends on how the change in state is carried out, namely, the heat change. However, in this instance we have specified how the change in state is to occur, namely, at constant volume. Therefore, if we measure the heat change at constant volume associated with a change in state, we have also measured the energy change.

Temperature is an especially important variable in biological systems. If the temperature is constant during a change in state, the process is *isothermal*. On the other hand, if the system is insulated so that no heat escapes or enters the system during the change in state ($q = 0$), the process is *adiabatic*.

1.6 ENTHALPY

Most experiments in the laboratory and in biological systems are done at constant pressure, rather than at constant volume. At constant pressure,

$$\Delta E = q_P - P(V_2 - V_1) \quad (1-17)$$

or

$$E_2 - E_1 = q_P - P(V_2 - V_1) \quad (1-18)$$

The heat change at constant pressure can be written as

$$q_P = (E_2 + PV_2) - (E_1 + PV_1) \quad (1-19)$$

This relationship can be simplified by defining a new state function, the *enthalpy*, H :

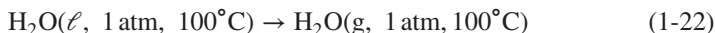
$$H = E + PV \quad (1-20)$$

The enthalpy is obviously a state function since E , P , and V are state functions. The heat change at constant pressure is then equal to the enthalpy change:

$$q_P = \Delta H = H_2 - H_1 \quad (1-21)$$

For biological reactions and processes, we will usually be interested in the enthalpy change rather than the energy change. It can be measured experimentally by determining the heat change at constant pressure.

As a simple example of how energy and enthalpy can be calculated, let us consider the conversion of liquid water to steam at 100°C and 1 atm, that is, boiling water:



The heat required for this process, $\Delta H (=q_P)$, is 9.71 kilocalories/mol. What is ΔE for this process? This can be calculated as follows:

$$\Delta E = \Delta H - \Delta(PV) = \Delta H - P\Delta V$$

$$\Delta V = V_g - V_\ell = 22.4 \text{ liters/mol} - 18.0 \times 10^{-3} \text{ liters/mol} \approx PV_g \approx RT$$

$$\Delta E = \Delta H - RT = 9710 - 2(373) = 8970 \text{ calories/mol}$$

Note that the Kelvin temperature must be used in thermodynamic calculations and that ΔH is significantly greater than ΔE .

Let us do a similar calculation for the melting of ice into liquid water



In this case, the measured heat change, $\Delta H (=q_p)$, is 1.44 kcal/mol. The calculation of ΔE parallels the previous calculation.

$$\Delta E = \Delta H - P\Delta V$$

$$\Delta V = V_\ell - V_g \approx 18.0 \text{ ml/mol} - 19.6 \text{ ml/mol} \approx -1.6 \text{ ml/mol}$$

$$P\Delta V \approx -1.6 \text{ ml atm} = -0.04 \text{ cal}$$

$$\Delta E = 1440 + 0.04 = 1440 \text{ cal/mol}$$

In this case, ΔE and ΔH are essentially the same. In general, they do not differ greatly in condensed media, but the differences can be substantial in the gas phase.

The two most common units for energy are the calorie and the joule (J). (One calorie equals 4.184 J.) The official MKS unit is the joule, but many research publications use the calorie. We will use both in this text, in order to familiarize the student with both units.

1.7 STANDARD STATES

Only changes in energy states can be measured. Therefore, it is arbitrary what we set as the zero for the energy scale. As a matter of convenience, a common zero has been set for both the energy and the enthalpy. Elements in their stablest forms at 25°C (298 K) and 1 atm are assigned an enthalpy of zero. This is called a *standard state* and is usually written as H_{298}^0 . The superscript means 1 atm and the subscript is the temperature in Kelvin.

As an example of how this concept is used, consider the formation of carbon tetrachloride from its elements:



$$\Delta H = H_{298(\text{CCl}_4)}^0 - H_{298(\text{C})}^0 - 2H_{298(\text{Cl}_2)}^0$$

$$\Delta H = H_{298(\text{CCl}_4)}^0 \quad (1-24)$$

The quantity $H_{298(\text{CCl}_4)}^0$ is called the heat of formation of carbon tetrachloride. Tables of heats of formation are available for hundreds of compounds and are useful in calculating the enthalpy changes associated with chemical reactions (cf. Refs. 4, 5).

In the case of substances of biological interest in solutions, the definitions of standard states and heats of formation are a bit more complex. In addition to pressure and temperature, other factors must be considered such as pH, salt concentration, metal ion concentration, etc. A universal definition has not been established. In practice, it is best to use heats of formation under a defined set of conditions and likewise to define the standard state as these conditions. Tables of heats of formation for some compounds of biological interest are given in Appendix 4 (6). A prime is often added to the symbol for these heats of formation ($H_f^{o'}$) to indicate the unusual nature of the

standard state. We will not make that distinction here, but it is essential that a consistent standard state is used when making thermodynamic calculations for biological systems.

A useful way of looking at chemical reactions is as algebraic equations. A characteristic enthalpy can be assigned to each product and reactant. Consider the “reaction”



For this reaction, $\Delta H = H_{\text{products}} - H_{\text{reactants}}$, or

$$\Delta H = dH_D + cH_C - aH_A - bH_B$$

where the H_i are molar enthalpies. At 298 K and 1 atm, the molar enthalpies of the elements are zero, whereas for compounds, the molar enthalpies are equal to the heats of formation, which are tabulated. Before we apply these considerations to biological reactions, a brief digression will be made to discuss how heats of reactions are determined experimentally.

1.8 CALORIMETRY

The area of science concerned with the measurement of heat changes associated with chemical reactions is designated as calorimetry. Only a brief introduction is given here, but it is important to relate the theoretical concepts to laboratory experiments. To begin this discussion, we will return to our earlier discussion of heat changes and the heat capacity, Eq. (1-3). Since the heat change depends on how the change in state is carried out, we must be more precise in defining the heat capacity. The two most common conditions are constant volume and constant pressure. The heat changes in these cases can be written as

$$dq_V = dE = C_V dT \quad (1-26)$$

$$dq_P = dH = C_P dT \quad (1-27)$$

A more exact mathematical treatment of these definitions would make use of partial derivatives, but we will avoid this complexity by using subscripts to indicate what is held constant. These equations can be integrated to give

$$\Delta E = \int_{T_1}^{T_2} C_V dT \quad (1-28)$$

$$\Delta H = \int_{T_1}^{T_2} C_P dT \quad (1-29)$$

Thus, heat changes can readily be measured if the heat capacity is known. The heat capacity of a substance can be determined by adding a known amount of heat to the

substance and determining the resulting increase in temperature. The known amount of heat is usually added electrically since this permits very precise measurement. (Recall that the electrical heat is I^2R , where I is the current and R is the resistance of the heating element.) If heat is added repeatedly in small increments over a large temperature range, the temperature dependence of the heat capacity can be determined. Tabulations of heat capacities are available and are usually presented with the temperature dependence described as a power series:

$$C_p = a + bT + cT^2 + \dots \quad (1-30)$$

where a , b , c , ... are constants determined by experiment.

For biological systems, two types of calorimetry are commonly carried out—batch calorimetry and scanning calorimetry. In batch calorimetry, the reactants are mixed together and the ensuing temperature rise (or decrease) is measured. A simple experimental setup is depicted in Figure 1-4, where the calorimeter is a Dewar flask and the temperature increase is measured by using a thermocouple or thermometer.

For example, if we wished to measure the heat change for the hydrolysis of adenosine 5'-triphosphate (ATP),



a solution of known ATP concentration would be put in the Dewar at a defined pH, metal ion concentration, buffer, etc. The reaction would be initiated by adding a small amount of adenosine triphosphatase (ATPase), an enzyme that efficiently catalyzes the hydrolysis, and the subsequent temperature rise is measured. The enthalpy of reaction can be calculated from the relationship

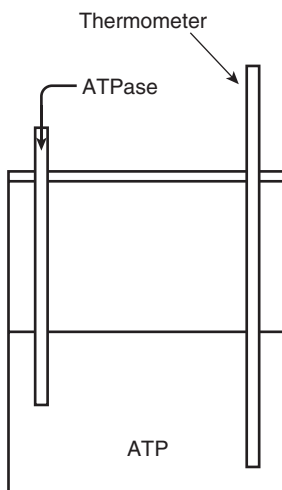


FIGURE 1-4. Schematic representation of a simple batch calorimeter. The insulated vessel is filled with a solution of ATP in a buffer containing salt and Mg^{2+} . The hydrolysis of ATP is initiated by the addition of the ATPase enzyme and the subsequent rise in temperature is measured.

$$\Delta H = C_p \Delta T \quad (1-32)$$

The heat capacity of the system is calculated by putting a known amount of heat into the system through an electrical heater and measuring the temperature rise of the system. The enthalpy change calculated is for the number of moles of ATP in the system. Usually, the experimental result is reported as a molar enthalpy, that is, the enthalpy change for a mole of ATP being hydrolyzed. This result can be obtained by dividing the observed enthalpy change by the moles of ATP hydrolyzed. Actual calorimeters are much more sophisticated than this primitive experimental setup. The calorimeter is well insulated, mixing is carried out very carefully, and very precise temperature measurements are made with a thermocouple. The enthalpy changes for many biological reactions have been measured, but unfortunately this information is not conveniently tabulated in a single source. However, many enthalpies of reaction can be derived from the heats of formation in the table in Appendix 4.

Scanning calorimetry is a quite different experiment and measures the heat capacity as a function of temperature. In these experiments, a known amount of heat is added to the system through electrical heating and the resulting temperature rise is measured. Very small amounts of heat are used, so the temperature changes are typically very small. This process is repeated automatically so that the temperature of the system slowly rises. The heat capacity of the system is calculated for each heat increment as $q_p/\Delta T$ and the data are presented as a plot of C_p versus T .

This method has been used, for example, to study protein unfolding and denaturation. Proteins unfold as the temperature is raised, and denaturation usually occurs over a very narrow temperature range. This is illustrated schematically in Figure 1-5, where the fraction of denatured protein, f_D , is plotted versus the temperature along with the corresponding plot of heat capacity, C_p , versus temperature.

As shown in Figure 1-5, the plot of heat capacity versus temperature is a smooth, slowly rising curve for the solvent. With the protein present, a peak in the curve occurs

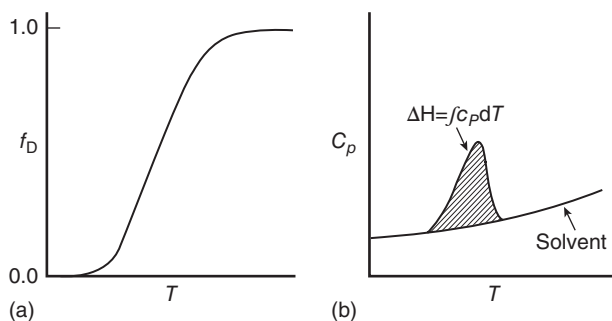


FIGURE 1-5. Schematic representation of the denaturation of a protein and the resulting change in heat capacity, C_p . In (a), the fraction of denatured protein, f_D , is shown as a function of temperature, T . In (b), the heat capacity, as measured by scanning calorimetry, is shown as a function of temperature. The lower curve is the heat capacity of the solvent. The hatched area is the excess heat capacity change due to the protein denaturing and is equal to ΔH for the unfolding.

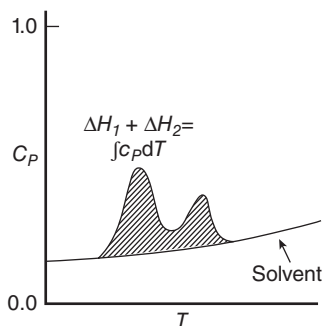


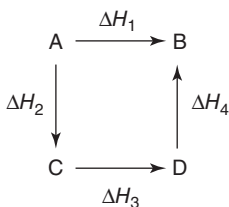
FIGURE 1-6. Schematic representation of a calorimeter scan in which the denaturation occurs in two steps. The hatched area permits the sum of the enthalpy changes to be determined, and the individual enthalpies of the unfolding reactions can be determined by a detailed analysis. As in Figure 1-5, C_p is the measured heat capacity and T is the temperature.

as the protein is denatured. The enthalpy change associated with denaturation is the area under the peak (striped area = $\int C_p dT$). In some cases, the protein denaturation may occur in multiple stages, in which case more than one peak can be seen in the heat capacity plot. This is shown schematically in Figure 1-6 for a two-stage unfolding process.

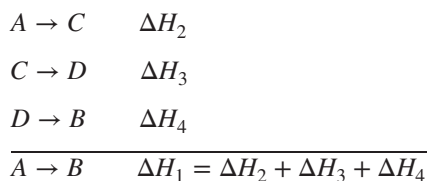
The enthalpies associated with protein unfolding are often interpreted in molecular terms such as hydrogen bonds, electrostatic interactions, and hydrophobic interactions. It should be borne in mind that these interpretations are not inherent in thermodynamic quantities, which do not explicitly give information at the molecular level. Consequently, such interpretations should be scrutinized very critically.

1.9 REACTION ENTHALPIES

We now return to a consideration of reaction enthalpies. Because the enthalpy is a state function, it can be added and subtracted for a sequence of reactions—it does not matter how the reaction occurs or in what order. In this regard, chemical reactions can be considered as algebraic equations. For example, consider the reaction cycle below:



If these reactions are written sequentially, it can readily be seen how the enthalpies are related.

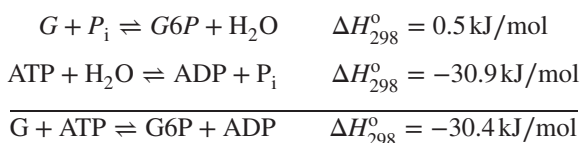


This ability to relate enthalpies of reaction in reaction cycles in an additive fashion is often called Hess's law, although it really is derived from thermodynamic principles as discussed. We will find that this "law" is extremely useful, as it allows determination of the enthalpy of reaction without studying a reaction directly if a sequence of reactions is known that can be added to give the desired reaction.

As an illustration, we will calculate the enthalpy of reaction for the transfer of a phosphoryl group from ATP to glucose, a very important physiological reaction catalyzed by the enzyme hexokinase.



The standard enthalpy changes for the hydrolysis of these four compounds are given in Table 1-1. These data are for very specific conditions: $T = 298 \text{ K}$, $P = 1 \text{ atm}$, $\text{pH} = 7.0$, $\text{pMg} = 3$, and an ionic strength of 0.25 M . The ionic strength is a measure of the salt concentration that takes into account the presence of both monovalent and divalent ions ($= \frac{1}{2} \sum c_i z_i^2$, where c_i is the concentration of each ion, z_i is its charge, and the sum is over all of the ions present). The enthalpy change for the hexokinase reaction can easily be calculated from these data:



The ability to calculate thermodynamic quantities for biochemical reactions that have not yet been studied is very useful. Even if data are not available to deal with the

TABLE 1-1. Standard Enthalpy Changes of Hydrolysis

Reaction	ΔH_{298}° (kJ/mol)
$\text{ATP} + \text{H}_2\text{O}(\ell) \rightleftharpoons \text{ADP} + P_i$	-30.9
$\text{ADP} + \text{H}_2\text{O}(\ell) \rightleftharpoons \text{AMP} + P_i$	-28.9
$\text{AMP} + \text{H}_2\text{O}(\ell) \rightleftharpoons \text{A} + P_i$	-1.2
$\text{G6P} + \text{H}_2\text{O}(\ell) \rightleftharpoons \text{G} + P_i$	-0.5

reaction of specific interest, very often data are available for closely related reactions. Appendix 5 contains a tabulation of ΔH_{298}° for some biochemical reactions.

The enthalpy change associated with the hexokinase reaction could also be derived from the heats of formation in the table in Appendix 4:

$$\Delta H = H_{f,ADP}^{\circ} + H_{f,G6P}^{\circ} - H_{f,ATP}^{\circ} - H_{f,G}^{\circ}$$

$$\Delta H = -2000.2 - 2279.1 + 2981.8 + 1267.1 = -30.4 \text{ kJ/mol}$$

In point of fact, the heats of formation are usually derived from measured heats of reaction as these are the primary experimental data.

A source of potential confusion is the practice of reporting enthalpies of reaction as “per mole.” There is no ambiguity for the hexokinase reaction as written above. However, in many cases, the stoichiometric coefficients for reactants and products differ. For example, the reaction catalyzed by the enzyme myokinase is



Even though 2 moles of ADP are used, the reaction enthalpy is referred to as “per mole.” The reaction enthalpy is always given as “per mole of reaction as it is written.”

It is important, therefore, that the equation for the reaction under consideration be explicitly stated. The myokinase reaction could be written as



In this case, the reaction enthalpy per mole would be one-half of that reported for Eq. (1-34).

1.10 TEMPERATURE DEPENDENCE OF THE REACTION ENTHALPY

In principle, the enthalpy changes as the pressure and temperature change. We will not worry about the dependence of the enthalpy on pressure, as it is usually very small for reactions in condensed phases. The temperature dependence of the enthalpy is given by Eq. (1-27). This can be used directly to determine the temperature dependence of reaction enthalpies. If we assume the standard state enthalpy is known for each reactant, then the temperature dependence of the enthalpy for each reactant, i , is

$$H_{T,i} = H_{298,i}^{\circ} + \int_{298}^T C_{P,i} dT \quad (1-36)$$

If we apply this relationship to the reaction enthalpy for the generalized reaction of Eq. (1-25), we obtain the following:

$$\Delta H_T = cH_{T,C} + dH_{T,D} - aH_{T,A} - bH_{T,B}$$

$$\Delta H_T = \Delta H_{298}^{\circ} + \int_{298}^T \Delta C_P dT$$

with

$$\Delta H_{298}^{\circ} = cH_{298,C}^{\circ} + dH_{298,D}^{\circ} - aH_{298,A}^{\circ} - bH_{298,B}^{\circ}$$

and

$$\Delta C_P = cC_{P,C} + dC_{P,D} - aC_{P,A} - bC_{P,B}$$

More generally,

$$\Delta H_T = \Delta H_{T_0} + \int_{T_0}^T \Delta C_P dT \quad (1-37)$$

Equation (1-37) is known as Kirchhoff's law. It can also be stated in differential form:

$$d\Delta H/dT = \Delta C_P \quad (1-38)$$

It is important to remember that this discussion of the temperature dependence of the reaction enthalpy assumes that the pressure is constant.

The conclusion of these considerations of reaction enthalpies is that available tabulations are often sufficient to calculate the reaction enthalpy of many biological reactions. Moreover, if this is carried out at a standard temperature, the reaction enthalpy at other temperatures can be calculated if appropriate information about the heat capacities is known or estimated. For most chemical reactions of biological interest, the temperature dependence of the reaction enthalpy is small. In contrast, for processes such as protein folding and unfolding, the temperature dependence is often significant and must be taken into account in data analysis and thermodynamic calculations. This will be discussed further in Chapter 3.

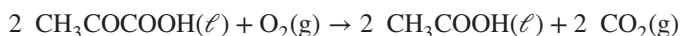
The first law of thermodynamics, namely, the definition of energy and its conservation, is obviously of great importance in understanding the nature of chemical reactions. As we shall see, however, the first law is not sufficient to understand what determines chemical equilibria.

REFERENCES

1. P. Atkins and J. de Paul, *Physical Chemistry* (9th edition), W. H. Freeman and Co., New York, NY, 2010.
2. I. Tinoco, Jr., K. Sauer, & J. C. Wang, J. D. Puglisi, G. Harbison, and D. Rovnyak, *Physical Chemistry: Principles and Applications to the Biological Sciences*, 5th edition, Prentice Hall, Englewood Cliffs, NJ, 2013.
3. D. Eisenberg and D. Crothers, *Physical Chemistry with Applications to the Life Sciences*, Benjamin/Cummings, Menlo Park, CA, 1979.
4. *The NBS Tables of Thermodynamic Properties*, D. D. Wagman et al. (eds.), *J. Phys. Chem. Ref. Data*, 11, Suppl. 2, 1982.
5. D. R. Stull, E. E. Westrum, Jr., and G. C. Sinke, *The Chemical Thermodynamics of Organic Compounds*, Wiley, New York, 1969.
6. R. A. Alberty, *Arch. Biochem. Biophys.* 353, 116 (1998).

PROBLEMS

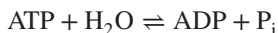
- 1-1.** When a gas expands rapidly through a valve, you often feel the valve get colder. This is an adiabatic expansion ($q = 0$). Calculate the decrease in temperature of 1.0 mol of ideal gas as it is expanded from 0.20 to 1.00 liters under the conditions given subsequently. Assume a constant volume molar heat capacity, C_V , of $\frac{3}{2}R$. Note that the energy, E , of an ideal gas depends only on the temperature: it is independent of the volume of the system.
- The expansion is irreversible with an external pressure of 1 atm and an initial temperature of 300 K.
 - The expansion is reversible with an initial temperature of 300 K.
 - Calculate ΔE for the changes in state described in parts A and B.
 - Assume that the expansion is carried out *isothermally* at 300 K, rather than adiabatically. Calculate the work done if the expansion is carried out irreversibly with an external pressure of 1.0 atm.
 - Calculate the work done if the isothermal expansion is carried out reversibly.
 - Calculate q and ΔE for the changes in state described in parts D and E.
- 1-2. a.** Calculate the enthalpy change for the conversion of glucose [$\text{C}_6\text{H}_{12}\text{O}_6(\text{s})$] and oxygen [$\text{O}_2(\text{g})$] to $\text{CO}_2(\text{aq})$ and $\text{H}_2\text{O}(\ell)$ under standard conditions. The standard enthalpies of formation of glucose(s), $\text{CO}_2(\text{aq})$, and $\text{H}_2\text{O}(\ell)$ are -304.3 , -98.7 , and -68.3 kcal/mol, respectively.
- b.** When organisms metabolize glucose, approximately 50% of the energy available is utilized for chemical and mechanical work. Assume that 25% of the total energy from eating 1 mole of glucose can be utilized to climb a mountain. How high a mountain can a 70 kg person climb?
- 1-3.** Calculate the enthalpy change for the oxidation of pyruvic acid to acetic acid under standard conditions.



The heats of combustion of pyruvic acid and acetic acid under standard conditions are -279 kcal/mol and -207 kcal/mol, respectively. Heats of combustion are determined by reacting pyruvic or acetic acid with $\text{O}_2(\text{g})$ to give $\text{H}_2\text{O}(\ell)$ and $\text{CO}_2(\text{g})$. *Hint:* First write balanced chemical equations for the combustion processes.

- 1-4.** Calculate the amount of water (in liters) that would have to be vaporized at 40°C (approximately body temperature) to expend the 2.5×10^6 calories of heat generated by a person in one day (commonly called sweating). The heat of vaporization of water at this temperature is 574 cal/g. We normally do not sweat that much. What is wrong with this calculation? If 1% of the energy produced as heat could be utilized as mechanical work, how large a weight could be lifted 1 meter?

- 1-5. a.** One hundred milliliters of 0.200 M ATP is mixed with an ATPase in a Dewar at 298 K, 1 atm, pH 7.0, pMg 3.0, and 0.25 M ionic strength. The temperature of the solution increases 1.48 K. What is ΔH° for the hydrolysis of ATP to adenosine 5'-diphosphate (ADP) and phosphate? Assume that the heat capacity of the system is 418 J/K.
- b.** The hydrolysis reaction can be written as



Under the same conditions, the hydrolysis of ADP



has a heat of reaction, ΔH° , of -28.9 kJ/mol. Under the same conditions, calculate ΔH° for the adenylate kinase reaction:



- 1-6.** The alcohol dehydrogenase reaction



removes ethanol from the blood. Use the enthalpies of formation in Appendix 4 to calculate ΔH° for this reaction. If 10.0 g of ethanol (a generous martini) is completely converted to acetaldehyde by this reaction, how much heat is produced or consumed?

Entropy and Gibbs Energy

2.1 INTRODUCTION

At the outset, we indicated that the primary objective of our discussion of thermodynamics is to understand chemical equilibrium in thermodynamic terms. On the basis of our discussion thus far, one possible conclusion is that chemical equilibria are governed by energy considerations and that the system will always proceed to the lowest energy state. This idea can be discarded quite quickly, as we know that some spontaneous reactions produce heat and some require heat. For example, the hydrolysis of ATP releases heat, $\Delta H_{298}^{\circ} = -30.9$ kJ/mol, whereas ATP and AMP are formed when ADP is mixed with myokinase, yet $\Delta H_{298}^{\circ} = +2.0$ kJ/mol under identical conditions. The conversion of liquids to gases requires heat, that is, ΔH is positive, even at temperatures above the boiling point. Clearly, the lowest energy state is not necessarily the most stable state.

What factor is missing? (At this point, traditional treatments of thermodynamics launch into a discussion of heat engines, a topic we will avoid.) The missing ingredient is the consideration of the probability of a given state. As a very simple illustration, consider three balls of equal size that are numbered 1, 2, and 3. These balls can be arranged sequentially in six different ways:

123 132 213 231 312 321

The energy state of all of these arrangements is the same, yet it is obvious that the probability of the balls being in sequence (123) is 1/6, whereas the probability of the balls being out of sequence is 5/6. In other words, the probability of a disordered state is much greater than the probability of an ordered state because a larger number of arrangements of the balls exists in the disordered state.

Molecular examples of this phenomenon can readily be found. A gas expands spontaneously into a vacuum even though the energy state of the gas does not change. This occurs because the larger volume has more positions available for molecules, so a greater number of arrangements, or more technically *microstates*, of molecules are possible. Clearly, probability considerations are not sufficient by

themselves. If this were the case, the stable state of matter would always be a gas. We know that solids and liquids are stable under appropriate conditions because they are energetically favored; that is, interactions between atoms and molecules result in a lower energy state. The real situation must involve a balance between energy and probability. This is a qualitative statement of what determines the equilibrium state of a system, but we will be able to be much more quantitative than this.

The second law states that disordered states are more probable than ordered states. This is done by defining a new state function, *entropy*, which is a measure of the disorder (or probability) of a state. Thermodynamics does not require this interpretation of the entropy, which is quasi-molecular. However, this is a much more intuitive way of understanding entropy than utilizing the traditional concept of heat engines. The more disordered a state, or the larger the number of available microstates, the higher the entropy. We already can see a glimmer of how the equilibrium state might be determined. At constant entropy, the energy should be minimized, whereas at constant energy, the entropy should be maximized. We will return to this topic a little later. First, we will define the entropy quantitatively.

2.2 STATEMENT OF THE SECOND LAW

A more formal statement of the second law is to define a new state function, the entropy, S , by the equation

$$dS = \frac{dq_{\text{rev}}}{T} \quad (2-1)$$

or

$$\Delta S = \int \frac{dq_{\text{rev}}}{T} \quad (2-2)$$

The temperature scale in this definition is Kelvin. This definition is not as straightforward as that for the energy. Note that this definition requires a reversible heat change, q_{rev} or dq_{rev} , yet entropy is a state function. At first glance, this seems quite paradoxical. The meaning of this is that the entropy change must be calculated by finding a reversible path. However, all reversible paths give the same entropy change, and the calculated entropy change is correct even if the actual change in state is carried out irreversibly. Although this appears to be somewhat confusing, consideration of some examples will help in understanding this concept.

The second law also includes important considerations about entropy: For a reversible change in state, the entropy of the universe is constant, whereas for an irreversible change in state, the entropy of the universe increases. Again, the second law cannot be proved, but we can demonstrate that without this law, events could transpire that are contrary to our everyday experience. Two examples are given in the following sections.

Without the second law, a gas could spontaneously compress! Let us illustrate this by considering the isothermal expansion of an ideal gas, V_1 to V_2 with constant T .

The entropy change is

$$\Delta S = \int \left(\frac{dq_{\text{rev}}}{T} \right) = \left(\frac{1}{T} \right) \int dq_{\text{rev}} = \frac{q_{\text{rev}}}{T} \quad (2-3)$$

For the isothermal expansion of an ideal gas, $\Delta E = 0$. (Because of the definition of an ideal gas, the energy, E , is determined by the temperature only and does not depend on the volume and pressure.) Therefore,

$$q_{\text{rev}} = -w_{\text{rev}} = \int_{V_1}^{V_2} P dV = \int_{V_1}^{V_2} \left(\frac{nRT}{V} \right) dV = nRT \ln \left(\frac{V_2}{V_1} \right) \quad (2-4)$$

and

$$\Delta S = nR \ln \left(\frac{V_2}{V_1} \right) \quad (2-5)$$

For an expansion, $V_2 > V_1$ and $\Delta S > 0$. For a compression, $V_2 < V_1$ and $\Delta S < 0$. The second law does not prohibit this situation, as we are considering the entropy change for the system, not the universe. This result is in accord with the intuitive interpretation of entropy previously discussed: A larger volume has more positions for the gas to occupy and consequently a higher entropy.

We must now consider what happens to the entropy of the surroundings. For a reversible change, $\Delta S = -q_{\text{rev}}/T$, and the entropy of the universe is the sum of the entropy change for the system and that for the surroundings: $\Delta S = q_{\text{rev}}/T - q_{\text{rev}}/T = 0$, which is consistent with the second law. However, for an irreversible change the situation is different. Let us make this irreversible change by setting the external pressure equal to zero during the change in state. In that case, $w = 0$ and since $\Delta E = 0$, $q = 0$. Therefore, no heat is lost by the surroundings. The entropy change for the universe is

$$\Delta S = \Delta S_{\text{gas}} + \Delta S_{\text{surr}} = nR \ln \left(\frac{V_2}{V_1} \right) + 0$$

The second law says that the entropy change of the universe must be greater than zero, which requires that $V_2 > V_1$. In other words, a spontaneous compression cannot occur. This is not required by the first law.

As a second example, consider two blocks at different temperatures, T_h and T_c , where h and c designate hot and cold so $T_h > T_c$. We will put the blocks together so that heat is transferred from the hot block to the cold block. The entropy changes in the two blocks are given by

$$dS_c = \frac{dq}{T_c} \quad \text{and} \quad dS_h = -\frac{dq}{T_h}$$

If the two blocks are considered to be the universe, the entropy change of the universe is

$$dS_c + dS_h = dq \left(\frac{1}{T_c} - \frac{1}{T_h} \right) > 0$$

As predicted by the second law, the entropy of the universe increases. What if the heat flows from the cold block to the hot block? Then the sign of the heat change is reversed and

$$dS_c + dS_h = dq \left(\frac{1}{T_h} - \frac{1}{T_c} \right) < 0$$

This predicts that the entropy of the universe would decrease, which is impossible according to the second law. Thus, heat cannot flow from the cold bar to the hot bar. This is not prohibited by the first law.

Exceptions to the second law can be used to create perpetual motion machines, which we know are impossible. These are sometimes called perpetual motion machines of the second kind, whereas perpetual motion machines created by exceptions to the first law are called perpetual motion machines of the first kind.

2.3 CALCULATION OF THE ENTROPY

A reversible path must always be found to calculate the entropy change. At constant volume, the relationship

$$dq_{\text{rev}} = nC_V dT \quad (2-6)$$

can be used, whereas at constant pressure

$$dq_{\text{rev}} = nC_P dT \quad (2-7)$$

or

$$\Delta S = \int \frac{nC_P dT}{T} \quad (2-8)$$

Entropy changes for phase changes are particularly easy to calculate since they occur at constant temperature and pressure. At constant temperature and pressure,

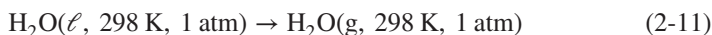
$$\Delta S = \frac{q_{\text{rev}}}{T} = \frac{\Delta H}{T} \quad (2-9)$$

For example, for the process

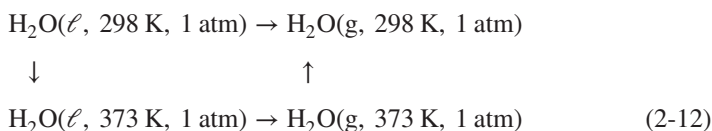


$\Delta H = 2380 \text{ cal/mol}$ and $\Delta S = 2380/279 = 8.53 \text{ cal/(mol K)} = 8.53 \text{ eu}$. Here 1 cal/(mol K) is defined as an entropy unit, eu. The entropy change for the reverse process is -8.53 eu . Note that the reverse process is not prohibited by the second law since it is the entropy change for the system, not the universe. Also note that, as expected, going from a solid to a liquid involves a positive entropy change since the liquid state is more disordered than the solid.

Although it is easy to state that the entropy change can be calculated for an irreversible process by finding a reversible way of going from the initial state to the final state, this is not always easy to do. This is usually not a matter of great consequence for our considerations, but we will consider one example to illustrate the process. Let us determine the entropy change for the following process:



This change is not a reversible change in state, as we know that the normal boiling point of water at 1 atm is 373 K. A possible reversible cycle that would go from the initial state to the final state at constant pressure is



The entropy change for the bottom process, which is reversible, is simply $\Delta H/T = 9710/373 = 26 \text{ cal}/(\text{mol K})$. The entropy change for the left-hand side of the square is (Eq. (2-8))

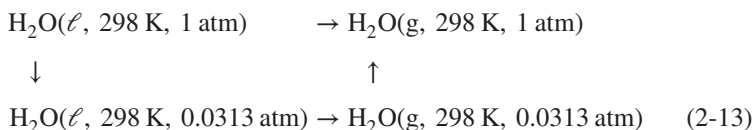
$$\Delta S = C_p \ln \left(\frac{T_2}{T_1} \right) = 18 \ln \left(\frac{373}{298} \right) = 4 \text{ cal}/(\text{mol K})$$

and for the right-hand side of the square is

$$\Delta S = C_p \ln \left(\frac{T_2}{T_1} \right) = 8.0 \ln \left(\frac{298}{373} \right) = -1.8 \text{ cal}/(\text{mol K})$$

(The heat capacity of $\text{H}_2\text{O}(\text{g})$ from 298 to 373 K is assumed to be that of $\text{H}_2\text{O}(\text{g})$ at 373 K.) The entropy changes for these three reversible processes can be added to give the entropy change for the change in state given in Eq. (2-11): $28 \text{ cal}/(\text{mol K})$.

An alternative reversible path that can be constructed lowers the pressure to the equilibrium vapor pressure of water at 298 K. The corresponding constant temperature cycle is



In this case, we would have to calculate the change in entropy as the pressure is lowered and raised. This can easily be done but is beyond the scope of this presentation of thermodynamics. The point of this exercise is to illustrate how entropy changes can be calculated for irreversible as well as reversible processes, and multiple reversible processes can be found.

In principle, the entropy can be calculated from statistical considerations. Boltzmann derived a relationship between the entropy, S , and the number of microstates, W :

$$S = k_B \ln W \quad (2-14)$$

where k_B is Boltzmann's constant, 1.38×10^{-23} J/K. It is rarely possible to determine the number of microstates although the number of microstates could be calculated from Eq. (2-14) if the entropy is known. For a simple case, such as the three numbered balls with which we started our discussion of the second law, this calculation can readily be done. The system has $3!$ microstates and an entropy of 1.51×10^{-23} J/K. Any ordered sequence – for example 1, 2, 3 – has only one microstate, so $W = 1$, and $S = 0$. Since the disordered state has a higher entropy, this predicts that the balls will spontaneously disorder and that an ordered state is much less likely. The statistical concept of entropy is discussed more fully in Chapter 15.

2.4 THIRD LAW OF THERMODYNAMICS

We will not dwell on the third law as the details are of little consequence in biology. The important fact for us is that the third law establishes a zero for the entropy scale. Unlike the energy, entropy has an absolute scale. The third law can be stated as follows: the entropy of perfect crystals of all pure elements and compounds is zero at absolute zero. The tricky points of this law are the meanings of “perfect” and “pure,” but we will not discuss this in detail. It is worth noting that a perfect crystal has one microstate, and, therefore, an entropy of zero according to Eq. (2-14).

The absolute standard entropy can be determined from measurements of the temperature dependence of the heat capacity using the relationship

$$S_{298}^{\circ} = \int_0^{298} \frac{C_P dT}{T} \quad (2-15)$$

Here the entropy at absolute zero has been assumed to be zero in accordance with the third law. A plot of C_P/T versus T gives a curve such as that shown in Figure 2-1. The area under the curve is the absolute standard entropy. Tables of S_{298}° are readily available (cf. Refs. 4 and 5 in Chapter 1). The entropy at temperatures other than 298 K can be calculated from the relationship

$$S_T = S_{298}^{\circ} + \int_{298}^T \frac{C_P dT}{T} \quad (2-16)$$

For a chemical reaction,

$$\Delta S_T = \Delta S_{298}^{\circ} + \int_{298}^T \frac{\Delta C_P dT}{T} \quad (2-17)$$

These relationships are analogous to those used for the enthalpy. Indeed, entropies of reactions can be calculated in a similar manner to enthalpy changes.

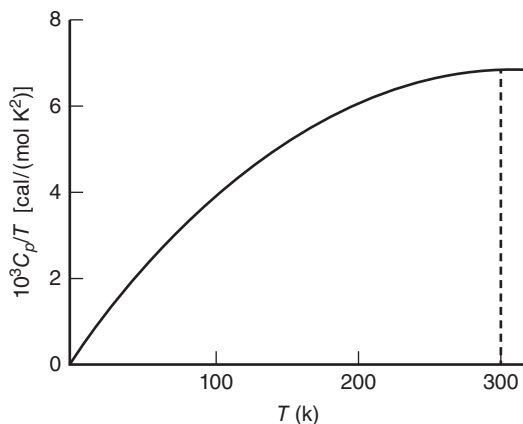


FIGURE 2-1. A plot of the constant-pressure heat capacity divided by the temperature, C_p/T , versus the temperature, T , for graphite. The absolute entropy of graphite at 300 K is the area under the curve up to the dashed line (Eq. (2-15)). If phase changes occur as the temperature is changed, the entropy changes associated with the phase changes must be added to the area under the curve of the C_p/T versus T plot.

2.5 MOLECULAR INTERPRETATION OF ENTROPY

We will now consider a few examples of absolute entropies and entropy changes for chemical reactions, and how they might be interpreted in molecular terms. The absolute entropies for water as a solid, liquid, and gas at 273 K and 1 atm are 41.0, 63.2, and 188.3 J/(mol K), respectively. The molecular interpretation of these numbers is straightforward, namely, that solid is more ordered than liquid, which is more ordered than gas; therefore, the solid has the lowest entropy and the gas the highest.

Standard entropy changes for some chemical reactions are given in Table 2-1. As expected, the entropy change is negative for the first two reactions in the gas phase as the number of moles of reactants is greater than the number of moles of products. Of course, these reactions could have been written in the opposite direction. The entropy change would then have the opposite sign. At first glance, the result for

TABLE 2-1. Standard Entropy Changes for Selected Reactions

Reaction	ΔS_{298}° [J/(mol K)]
$\text{H(g)} + \text{H(g)} \rightleftharpoons \text{H}_2(\text{g})$	-98.7
$2 \text{ H}_2(\text{g}) + \text{O}_2(\text{g}) \rightleftharpoons 2 \text{ H}_2\text{O(g)}$	-88.9
$\text{H}^+(\text{aq}) + \text{OH}^-(\text{aq}) \rightleftharpoons \text{H}_2\text{O}(\ell)$	+80.7
Cytidine 2' -monophosphate + Ribonuclease	-54 ^a
$\text{A} \rightleftharpoons \text{Enzyme} - \text{inhibitor complex}$	

^aH. Naghibi, A. Tamura, and J. M. Sturtevant, *Proc. Natl. Acad. Sci. USA* **92**, 5597 (1995).

the third reaction in the table is surprising, as the number of moles of reactants is greater than the number of moles of products. This results because the solvent must be included in the molecular interpretation of the observed entropy change. Ions in water interact strongly with water so that highly ordered water molecules exist around the ions.

When the neutral species is formed, these highly ordered water molecules become less ordered. Thus, the entropy change for water is very positive, much more positive than the expected entropy decrease for H^+ and OH^- when they form water. This simple example indicates that considerable care must be taken in interpreting entropy changes for chemical reactions in condensed media. The entropy of the entire system under consideration must be taken into account. The final entry is for the binding of a ligand to an enzyme. In this case, no reasonable interpretation of the entropy change is possible as three factors come into play: the loss in entropy as two reactants become a single entity, the changes in the structure of water, and structural changes in the protein. The fact that the entropy change is comparable to the value expected for the combination of two molecules to produce one molecule is simply fortuitous. Extreme caution should be taken in making molecular interpretations of thermodynamic changes in complex systems.

We have established all of the thermodynamic principles necessary to discuss chemical equilibrium. We will now apply these principles to develop a general framework for dealing with chemical reactions.

2.6 GIBBS ENERGY

In a sense, we have reached our goal. We have developed thermodynamic criteria for the occurrence of reversible and irreversible (spontaneous) processes, namely, for the universe, $\Delta S = 0$ for reversible processes and must be greater than zero for irreversible processes. Unfortunately, this is not terribly useful, as we are interested in what is happening in the system and require criteria that are easily applicable to chemical reactions. This can be achieved by defining a new thermodynamic state function, the Gibbs energy:

$$G = H - TS \quad (2-18)$$

(J. Willard Gibbs developed the science of thermodynamics virtually single-handedly around the turn of the century at Yale University. To this day, his collected works are prized possessions in the libraries of people seriously interested in thermodynamics.) In much of the literature and in some text books, G is called the free energy rather than the Gibbs energy. However, the official international nomenclature now defines G as the Gibbs energy.

At constant temperature,

$$\begin{aligned} \Delta G &= \Delta H - T\Delta S \\ &= q_P - T\Delta S = q_P - q_{\text{rev}} \end{aligned} \quad (2-19)$$

For a reversible process, $q_p = q_{\text{rev}}$, so $\Delta G = 0$. For irreversible processes, the situation is a bit more complex as the sign of the heat change must be considered. For an endothermic process, $q_p < q_{\text{rev}}$, so $\Delta G < 0$. For an exothermic process, the absolute value of q_p is greater than the absolute value of q_{rev} , but both are negative, so again, $\Delta G < 0$. This is the Gibbs energy change for the system and does not involve the surroundings. We now have developed criteria that tell us if a process occurs spontaneously at constant temperature and pressure. If $\Delta G < 0$, the change in state occurs spontaneously, whereas if $\Delta G > 0$, the reverse change in state occurs spontaneously. If $\Delta G = 0$, the system is at equilibrium. At constant volume and temperature, the spontaneity is determined by the Helmholtz energy, $A = E - TS$.

As with the energy and enthalpy, only differences in the Gibbs energy can be measured. Consequently, the zero of the Gibbs energy scale is arbitrary. As for the enthalpy, the zero of the scale is taken as the elements in their stable state at 1 atm and 298 K. Moreover, analogous to the enthalpy, tables of the Gibbs energies of formation of compounds are available so that Gibbs energy changes for chemical reactions can be calculated (cf. Refs. 4–6 in Chapter 1). Referring back to the reaction in Eq. (1-25),

$$\Delta G_{298}^{\circ} = cG_{298,C}^{\circ} + dG_{298,D}^{\circ} - aG_{298,A}^{\circ} - bG_{298,B}^{\circ} \quad (2-20)$$

where the Gibbs energies on the right-hand side of the equation are Gibbs energies per mole. We will discuss the temperature dependence of the Gibbs energy a bit later, but it is useful to remember that at constant temperature,

$$\Delta G = \Delta H - T\Delta S \quad (2-21)$$

Standard state Gibbs energy changes are available for biochemical reactions, although comprehensive tabulations do not exist, and the definition of “standard state” is more complex than in the gas phase as discussed previously. Standard state Gibbs energies of formation for some substances of biochemical interest are given in Appendix 4.

As an illustration of the concept of Gibbs energy, consider the conversion of liquid water to steam:



At the boiling point, $\Delta H = 9710 \text{ cal/mol}$ and $\Delta S = 26 \text{ eu}$. Therefore,

$$\Delta G = 9710 - 26T \quad (2-23)$$

At equilibrium, $\Delta G = 0$ and Eq. (2-23) gives $T = 373 \text{ K}$, the normal boiling point of water. If $T > 373 \text{ K}$, $\Delta G < 0$, and the change in state is spontaneous, whereas if $T < 373 \text{ K}$, $\Delta G > 0$, and the reverse process, condensation, occurs spontaneously.

Calculation of the Gibbs energy for a change in state is not always straightforward. As the entropy is part of the definition of the Gibbs energy, a reversible process must always be found to calculate the Gibbs energy change. The Gibbs energy change, of course, is the same regardless of how the change in state is accomplished since the Gibbs energy is a state function. As a simple example, again consider the change in

state in Eq. (2-11). This is not a reversible process since the two states are not in equilibrium. This is not a spontaneous change in state so that $\Delta G > 0$. In order to calculate the value of ΔG , we must think of a reversible path for carrying out the change in state. The two cycles in Eqs. (2-12) and (2-13) again can be used. In both cases, the bottom reaction is an equilibrium process, so $\Delta G = 0$. To calculate ΔG for the top reaction, we only need to add up the ΔG values for the vertical processes. These can be calculated from the temperature (upper cycle) and pressure (lower cycle) dependence of G . Such functional dependencies will be considered shortly. It will be a useful exercise for the reader to carry out the complete calculations.

2.7 CHEMICAL EQUILIBRIA

Although we now have developed criteria for deciding whether or not a process will occur spontaneously, they are not sufficient for consideration of chemical reactions. We know that chemical reactions are generally not “all or nothing” processes; instead, an equilibrium state is reached where both reactants and products are present. We will now derive a quantitative relationship between the Gibbs energy change and the concentrations of reactants and products. We will do this in detail for the simple case of ideal gases and by analogy for reactions in liquids.

The starting point for the derivation is the definition of Gibbs energy and its total derivative:

$$G = H - TS = E + PV - TS \quad (2-24)$$

$$dG = dE + PdV + VdP - TdS - SdT \quad (2-25)$$

Since $dE = dq + dw = TdS - PdV$ for a reversible process,

$$dG = VdP - SdT \quad (2-26)$$

(Although this relationship was derived for a reversible process, it is also valid for an irreversible process.) Let us now consider a chemical reaction of ideal gases at constant temperature. For one mole of each gas component,

$$dG = VdP = RTdP/P \quad (2-27)$$

We will reference all of our calculations to a pressure of 1 atm for each component. In thermodynamic terms, we have selected 1 atm as our *standard state*. If Eq. (2-27) is now integrated from $P_0 = 1$ atm to P ,

$$G = G^\circ + RT \ln \left(\frac{P}{P_0} \right) = G^\circ + RT \ln P \quad (2-28)$$

We will now return to our prototype reaction, Eq. (1-25), and calculate the Gibbs energy change. The partial pressures of the reactants are given in parentheses:

$$\begin{aligned}
 aA(P_A) + bB(P_B) &\rightleftharpoons cC(P_C) + dD(P_D) \\
 \Delta G &= cG_C + dG_D - aG_A - bG_B \\
 &= cG_C^\circ + dG_D^\circ - aG_A^\circ - bG_B^\circ + cRT \ln P_C \\
 &\quad + dRT \ln P_D - aRT \ln P_A - bRT \ln P_B
 \end{aligned} \tag{2-29}$$

or

$$\Delta G = \Delta G^\circ + RT \ln \left(\frac{P_C^c P_D^d}{P_A^a P_B^b} \right) \tag{2-30}$$

ΔG is the Gibbs energy for the reaction in Eq. (2-29) when the system is not at equilibrium. At equilibrium, at constant temperature and pressure, $\Delta G = 0$, and Eq. (2-30) becomes

$$\Delta G^\circ = -RT \ln \left(\frac{P_C^c P_D^d}{P_A^a P_B^b} \right)_e = -RT \ln K \tag{2-31}$$

Here the subscript e has been used to designate equilibrium and K is the equilibrium constant.

We now have a quantitative relationship between the partial pressures of the reactants and the standard Gibbs energy change, ΔG° . The standard Gibbs energy change is a constant at a given temperature and pressure but will vary as the temperature and pressure change. If $\Delta G^\circ < 0$, then $K > 1$, whereas if $\Delta G^\circ > 0$, $K < 1$. A common mistake is to confuse the Gibbs energy change with the standard Gibbs energy change. The Gibbs energy change is always equal to zero at equilibrium and can be calculated from Eq. (2-30) when not at equilibrium. The standard Gibbs energy change is a constant representing the hypothetical reaction with all of the reactants and products at a pressure of 1 atm. It is equal to zero only if the equilibrium constant fortuitously is 1.

Biological reactions do not occur in the gas phase. What about Gibbs energy in solutions? Conceptually there is no difference. The molar Gibbs energy at constant temperature and pressure can be written as

$$G = G^\circ + RT \ln \left(\frac{c}{c_0} \right) \tag{2-32}$$

where c is the concentration and c_0 is the standard state concentration. A more correct treatment would define the molar free energy as

$$G = G^\circ + RT \ln a \tag{2-33}$$

where a is the thermodynamic activity and is dimensionless. However, the thermodynamic activity can be written as a product of an activity coefficient and the concentration. The activity coefficient can be included in the standard Gibbs energy, G° , which

gives rise to Eq. (2-32). We need not worry about this as long as the solution conditions are clearly defined with respect to salt concentration, pH, etc. It is not of great concern because all of the aforementioned complications can be included in the standard Gibbs energy change since, in practice, the standard Gibbs energy is determined by measuring the equilibrium constant under defined conditions. A more complete discussion of the concepts of standard states, activity, and activity coefficients is given in Chapter 4.

Finally, we should note that the Gibbs energy per mole at constant temperature and pressure is called the chemical potential, μ , in more sophisticated treatments of thermodynamics, but there is no need to introduce this terminology here. Chapter 4 provides a more comprehensive discussion of the chemical potential.

If we take the standard state as 1 mol/liter, then the results parallel to Eqs. (2-28), (2-30), and (2-31) are

$$G = G^\circ + RT \ln c \quad (2-34)$$

$$\Delta G = \Delta G^\circ + RT \ln \left(\frac{c_C^c c_D^d}{c_A^a c_B^b} \right) \quad (2-35)$$

$$\Delta G^\circ = -RT \ln \left(\frac{c_C^c c_D^d}{c_A^a c_B^b} \right)_e = -RT \ln K \quad (2-36)$$

Equations (2-34)–(2-36) summarize the thermodynamic relationships necessary to discuss chemical equilibria.

Note that, strictly speaking, the equilibrium constant is dimensionless as all of the concentrations are ratios, the actual concentration divided by the standard state concentration. However, practically speaking, it is preferable to report equilibrium constants with the dimensions implied by the ratio of concentrations in Eq. (2-36). The equilibrium constant is determined experimentally by measuring concentrations, and attributing dimensions to this constant assures that the correct ratio of concentrations is considered and that the standard state is precisely defined.

Consideration of the Gibbs energy also allows us to assess how the energy and entropy are balanced to achieve the final equilibrium state. Since $\Delta G = \Delta H - T\Delta S$ at constant T and P , it can be seen that a change in state is spontaneous if the enthalpy change is very negative and/or the entropy change is very positive. Even if the enthalpy change is unfavorable (positive), the change in state will be spontaneous if the $T\Delta S$ term is very positive. Similarly, even if the entropy change is unfavorable ($T\Delta S$ very negative), the change in state will be spontaneous if the enthalpy change is sufficiently negative. Thus, the final equilibrium achieved is a balance between the enthalpy (ΔH) and the entropy ($T\Delta S$).

2.8 PRESSURE AND TEMPERATURE DEPENDENCE OF THE GIBBS ENERGY

We will now return to the pressure and temperature dependence of the Gibbs energy. At constant temperature, the pressure dependence of the Gibbs energy follows directly from Eq. (2-26), namely,

$$dG = VdP \quad (2-37)$$

This equation can be integrated if the pressure dependence of the volume is known, as for an ideal gas. For a chemical reaction, Eq. (2-37) can be rewritten as

$$d\Delta G = \Delta VdP \quad (2-38)$$

where ΔV is the difference in volume between the products and reactants. The pressure dependence of the equilibrium constant at constant temperature follows directly,

$$d\Delta G^\circ = -RTd\ln K = \Delta VdP \quad (2-39)$$

or

$$\frac{d \ln K}{dP} = -\frac{\Delta V}{RT} \quad (2-40)$$

For most chemical reactions, the pressure dependence of the equilibrium constant is quite small so that it is not often considered in biological systems.

Equilibrium constants, however, frequently vary significantly with temperature. At constant pressure, Eq. (2-26) gives

$$dG = -SdT \quad (2-41)$$

or

$$d\Delta G = -\Delta SdT \quad (2-42)$$

Returning to the basic definition of free energy at constant temperature and pressure,

$$\Delta G = \Delta H - T\Delta S = \Delta H + T \left(\frac{d\Delta G}{dT} \right)$$

This equation can be divided by T^2 and rearranged as follows:

$$\begin{aligned} -\frac{\Delta G}{T^2} + \frac{\left(\frac{d\Delta G}{dT} \right)}{T} &= -\frac{\Delta H}{T^2} \\ \frac{d(\Delta G/T)}{dT} &= -\frac{\Delta H}{T^2} \end{aligned} \quad (2-43)$$

Equation (2-43) is an important thermodynamic relationship describing the temperature dependence of the Gibbs energy at constant pressure and is called the Gibbs-Helmholtz equation. The temperature dependence of the equilibrium constant follows directly:

$$\frac{d(\Delta G^\circ/T)}{dT} = -R \frac{d \ln K}{dT}$$

or

$$\frac{d \ln K}{dT} = \frac{\Delta H^\circ}{RT^2} \quad (2-44)$$

If ΔH° is independent of temperature, Eq. (2-44) can easily be integrated:

$$\begin{aligned} d \ln K &= \int_{T_1}^{T_2} \frac{\Delta H^\circ dT}{RT^2} \\ \ln \left(\frac{K_2}{K_1} \right) &= \frac{\Delta H^\circ}{R} \left(\frac{1}{T_1} - \frac{1}{T_2} \right) = \frac{(\Delta H^\circ/R)(T_2 - T_1)}{T_1 T_2} \end{aligned} \quad (2-45)$$

When carrying out calculations, the difference between reciprocal temperatures should never be used directly as it introduces a large error. Instead, the rearrangement in Eq. (2-45) should be used in which the difference between two temperatures occurs. With this equation and a knowledge of ΔH° , the equilibrium constant can be calculated at any temperature if it is known at one temperature.

What about the assumption that the standard enthalpy change is independent of temperature? This assumption is reasonable for many biological reactions if the temperature range is not too large. In some cases, the temperature dependence cannot be neglected and must be included explicitly in carrying out the integration of Eq. (2-44). The temperature dependence of the reaction enthalpy depends on the difference in heat capacities between the products and reactants as given by Eq. (1-38).

Examples of the temperature dependence of equilibrium constants are displayed in Figure 2-2. As predicted by Eq. (2-45), a plot of $\ln K$ versus $1/T$ is a straight line with a slope of $-\Delta H^\circ/R$. The data presented are for the binding of DNA to DNA-binding proteins (i.e., Zn fingers). The dissociation constants for binding are quite similar for both proteins at 22°C, 1.08×10^{-9} M (WT1 protein) and 3.58×10^{-9} M (EGR1 protein). However, as indicated by the data in the figure, the standard enthalpy changes are *opposite* in sign, +6.6 kcal/mol and -6.9 kcal/mol, respectively. Consequently, the standard entropy changes are also quite different, 63.3 eu and 15.3 eu, respectively. These data indicate that there are significant differences in the binding processes, despite the similar equilibrium constants at 22°C.

2.9 PHASE CHANGES

The criterion that $\Delta G = 0$ at equilibrium at constant temperature and pressure is quite general. For chemical reactions, this means that the Gibbs energy of the products is equal to the Gibbs energy of the reactants. For phase changes of pure substances, this

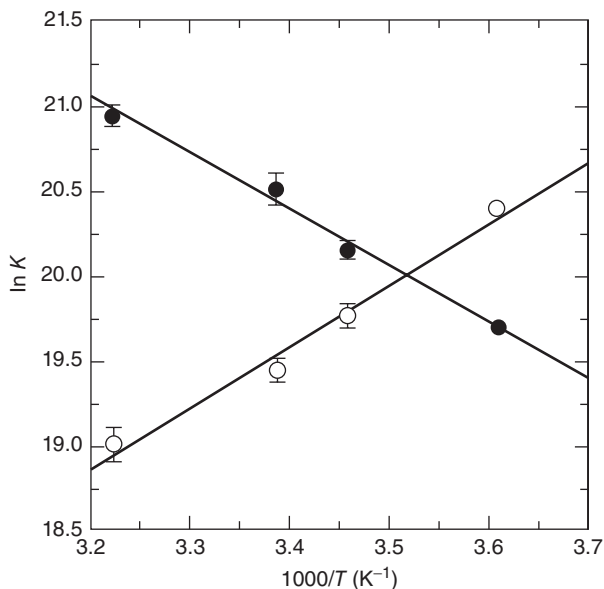


FIGURE 2-2. Temperature dependence of the equilibrium binding constant, K , for the binding of DNA to binding proteins (“zinc fingers”), WT1 (●) and EGR1 (○), to DNA. Adapted with permission from T. Hamilton, F. Borel, and P. J. Romaniuk, *Biochemistry* **37**, 2051 (1998). © 1998 American Chemical Society.

means that at equilibrium the Gibbs energies of the phases are equal. If we assume two phases, A and B, Eq. (2-26) gives

$$dG_A = V_A dP - S_A dT = dG_B = V_B dP - S_B dT$$

Rearrangement gives

$$\frac{dP}{dT} = \frac{\Delta S_{BA}}{\Delta V_{BA}} \quad (2-46)$$

where $\Delta S_{BA} = S_B - S_A$ and $\Delta V_{BA} = V_B - V_A$. (All of these quantities are assumed to be per mole for simplicity.) This equation is often referred to as the Clapeyron equation. Note that the entropy change can be written as

$$\Delta S_{BA} = \frac{\Delta H_{BA}}{T} \quad (2-47)$$

Equation (2-46) gives the slope of *phase diagrams*, plots of P versus T that are useful summaries of the phase behavior of a pure substance. As an example, the phase diagram of water is given in Figure 2-3.

The lines in Figure 2-3 indicate when two phases are in equilibrium and coexist. Only one phase exists in the open areas, and when the three lines meet, three phases are in equilibrium. This is called the triple point and can only occur at a pressure of

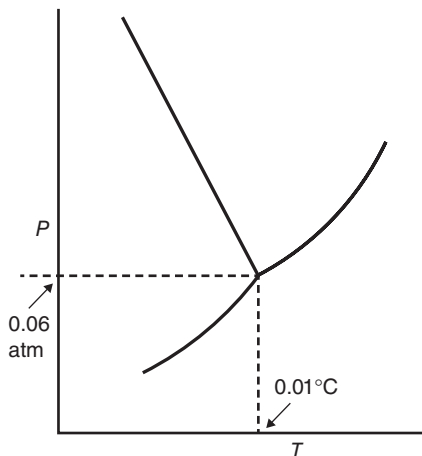


FIGURE 2-3. Schematic representation of the phase diagram of water with pressure, P , and temperature, T , as variables. The phase diagram is not to scale and is incomplete, as several different phases of solid water are known. If volume is included as a variable, a three-dimensional phase diagram can be constructed.

0.006 atm and 0.01°C . The slopes of the lines are given by Eq. (2-46). The number of phases that can coexist is governed by the phase rule,

$$f = c - p + 2 \quad (2-48)$$

where f is the number of degrees of freedom, c is the number of components, and p is the number of phases. This rule can be derived from the basic laws of thermodynamics. For a pure substance, the number of components is 1, so $f = 3 - p$. In the open spaces, one phase exists, so that the number of degrees of freedom is 2, which means that both T and P can be varied. Two phases coexist on the lines, which gives $f = 1$. Because only one degree of freedom exists, at a given value of P , T is fixed and vice versa. At the triple point, three phases coexist, and there are no degrees of freedom, which means that both P and T are fixed.

The study of phase diagrams is an important aspect of thermodynamics, and some interesting applications exist in biology. For example, membranes contain phospholipids. If phospholipids are mixed with water, they spontaneously form bilayer vesicles with the head groups pointing outward toward the solvent so that the interior of the bilayer is comprised of the hydrocarbon chains of the phospholipids. A schematic representation of the bilayer is given in Figure 2-4. The hydrocarbon chain can be either disordered, at high temperatures, or ordered, at low temperatures, as indicated schematically in the figure. The transformation from one form to the other behaves as a phase transition so that phase diagrams can be constructed for phospholipids. Moreover, the phase diagrams depend not only on the temperature but also on the phospholipid composition of the bilayer. This phase transition has biological implications in that the fluidity of the hydrocarbon portion of the membrane strongly

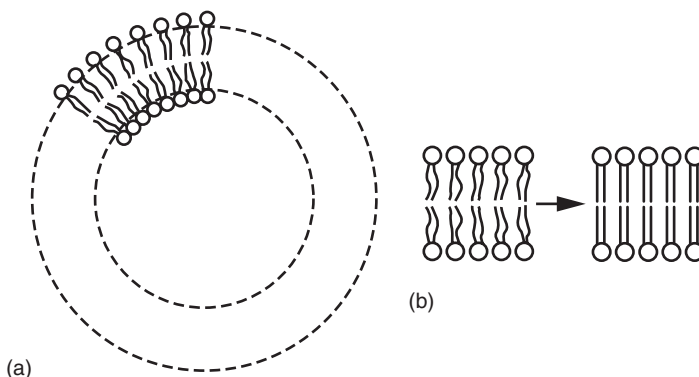


FIGURE 2-4. Schematic representation of the vesicles of phospholipids that are formed when phospholipids are suspended in water (a). Here the small circles represent the head groups and the wiggly lines the two hydrocarbon chains. A phase change can occur in the side chains from a disordered to an ordered state as sketched in (b).

affects the transport and mechanical properties of the membranes. Phase diagrams for phospholipids and phospholipid–water mixtures can be constructed by a variety of methods, including scanning calorimetry and nuclear magnetic resonance.

2.10 ADDITIONS TO THE GIBBS ENERGY

As a final consideration in developing the concept of Gibbs energy, we will return to our original development of equilibrium, namely, Eq. (2-33). In arriving at this expression for the molar Gibbs energy at constant temperature and pressure, the assumption was made that only $P - V$ work occurs. This is not true in general and for many biological systems in particular. If we went all the way back to the beginning of our development of the Gibbs energy changes associated with chemical equilibria (Eq. (2-26)) and derived the molar Gibbs energy at constant temperature and pressure, we would find that

$$G = G^{\circ} + RT \ln a + w_{\max} \quad (2-49)$$

where $w_{\max}$ is the maximum (reversible) non- $P - V$ work. For example, for an ion of charge z in a potential field Ψ , $w_{\max}$ is $zF\Psi$, where F is the Faraday. We have used this relationship in Chapter 1 to calculate the work of moving an ion across a membrane with an imposed voltage. In terms of transport of an ion across a membrane, this assumes that the ion is being transported from the inside to the outside with an external membrane potential of Ψ . This *extended chemical potential* is useful for discussing ion transport across membranes, as we shall see later. It could also be used for considering molecules in a gravitational or centrifugal field. (See Ref. 2 in Chapter 1 for a detailed discussion of this concept.)

We are now ready to consider some applications of thermodynamics to biological systems in more detail.

PROBLEMS

- 2-1.** One mole of an ideal gas initially at 300 K is expanded from 0.2 to 1.0 liter. Calculate ΔS for this change in state if it is carried out under the following conditions. Assume a constant volume heat capacity for the gas, C_V , of $\frac{3}{2} R$.
- The expansion is *reversible* and *isothermal*.
 - The expansion is *irreversible* and *isothermal*.
 - The expansion is *reversible* and *adiabatic* ($q = 0$).
 - The expansion is *irreversible* with an external pressure of 1 atm and *adiabatic*.
- 2-2.** The alcohol dehydrogenase reaction, which removes ethanol from your blood, proceeds according to the following reaction:



Under standard conditions (298 K, 1 atm, pH 7.0, pMg 3, and an ionic strength of 0.25 M), the standard enthalpies and Gibbs energies of formation of the reactants are as follows:

	H° (kJ/mol)	G° (kJ/mol)
NAD ⁺	−10.3	1059.1
NADH	−41.4	1120.1
Ethanol	−290.8	63.0
Acetaldehyde	−213.6	24.1

- Calculate ΔG° , ΔH° , and ΔS° for the alcohol dehydrogenase reaction under standard conditions.
 - Under standard conditions, what is the equilibrium constant for this reaction? Will the equilibrium constant increase or decrease as the temperature is increased?
- 2-3.** The equilibrium constant under standard conditions (1 atm, 298 K, pH 7.0) for the reaction catalyzed by fumarase



is 4.00. At 310 K, the equilibrium constant is 8.00.

- What is the standard Gibbs energy change, ΔG° , for this reaction?
- What is the standard enthalpy change, ΔH° , for this reaction? Assume that the standard enthalpy change is independent of temperature.
- What is the standard entropy change, ΔS° , at 298 K for this reaction?

- d. If the concentration of both reactants is equal to 0.01 M, what is the Gibbs energy change at 298 K? As the reaction proceeds to equilibrium, will more fumarate or L-malate form?
- e. What is the Gibbs energy change for this reaction when equilibrium is reached?

2-4. The following reaction is catalyzed by the enzyme creatine kinase:

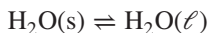


- a. Under standard conditions (1 atm, 298 K, pH 7.0, pMg 3.0, and an ionic strength of 0.25 M) with the concentrations of all reactants equal to 10 mM, the Gibbs energy change, ΔG , for this reaction is 13.3 kJ/mol. What is the standard Gibbs energy change, ΔG° , for this reaction?
- b. What is the equilibrium constant for this reaction?
- c. The standard enthalpies of formation for the reactants are as follows:

Creatine	−540 kJ/mol
Creatine phosphate	−1510 kJ/mol
ATP	−2982 kJ/mol
ADP	−2000 kJ/mol

What is the standard enthalpy change, ΔH° , for this reaction?

- d. What is the standard entropy change, ΔS° , for this reaction?
- 2-5. a. It has been proposed that the reason ice skating works so well is that the pressure from the blades of the skates melts the ice. Consider this proposal from the viewpoint of phase equilibria. The phase change in question is



Assume that ΔH for this process is independent of temperature and pressure and is equal to 80 cal/g. The change in volume, ΔV , is about -9.1×10^{-5} l/g. The pressure exerted is the force per unit area. For a 180 lb person and an area for the skate blades of about 6 in.², the pressure is 30 lb/in.² or about 2 atm. With this information, calculate the decrease in the melting temperature of ice caused by the skate blades. (Note that 1 cal = 0.04129 l atm.) Is this a good explanation for why ice skating works?

- b. A more efficient way of melting ice is to add an inert compound such as urea. (We will avoid salt to save our cars.) The extent to which the freezing point is lowered can be calculated by noting that the molar Gibbs energy of water must be the same in the solid ice and the urea solution. The molar Gibbs energy of water in the urea solution can be approximated as $G_{\text{liquid}}^\circ + RT \ln X_{\text{water}}$, where X_{water} is the mole fraction of water in the solution. The

molar Gibbs energy of the solid can be written as G_{solid}° . Derive an expression for the change in the melting temperature of ice by equating the Gibbs energies in the two phases, differentiating the resulting equation with respect to temperature, integrating from a mole fraction of 1 (pure solvent) to the mole fraction of the solution, and noting that $\ln X_{\text{water}} = \ln(1 - X_{\text{urea}}) \approx -X_{\text{urea}}$. (This relationship is the series expansion of the logarithm for small values of X_{urea} . Since the concentration of water is about 55 M, this is a good approximation.) With the relationship derived, estimate the decrease in the melting temperature of ice for an 1 M urea solution. The heat of fusion of water is 1440 cal/mol.

- 2-6.** What is the maximum amount of work that can be obtained from hydrolyzing 1 mole of ATP to ADP and phosphate at 298 K? Assume that the concentrations of all reactants are 0.01 M and ΔG° is -32.5 kJ/mol. If the conversion of free energy to mechanical work is 100% efficient, how many moles of ATP would have to be hydrolyzed to lift a 100 kg weight 1 meter high?

Applications of Thermodynamics to Biological Systems

3.1 BIOCHEMICAL REACTIONS

From the discussion in the previous chapter, it should be clear that standard Gibbs energy and enthalpy changes can be calculated for chemical reactions from available tables of standard Gibbs energies and enthalpies of formation (e.g., Appendix 4). Even if a particular reaction has not been studied, it is frequently possible to combine the standard Gibbs energy and enthalpy changes of other known reactions to obtain the standard Gibbs energy change of the desired reaction (e.g., Appendix 5). Knowledge of the standard Gibbs energy change permits calculation of the equilibrium constant and vice versa. When considering biochemical reactions, the exact nature of the “standard state” can be confusing. Usually, the various possible ionization states of the reactants are not specified, so that the total concentration of each species is used in the expression for the equilibrium constant. For example, for ATP, the standard state would normally be 1 M, but this includes all possible states of protonation, and if magnesium ion is present, the sum of the metal complex and uncomplexed species. Furthermore, pH 7 is usually selected as the standard condition so that the activity of the hydrogen ion is set equal to 1 at pH 7 and the standard Gibbs energy of formation of the hydrogen ion is set equal to zero at pH 7. Therefore, the hydrogen ion is not usually included when writing stoichiometric equations. Finally, the activity of water is set equal to 1 for reactions in dilute solutions. This assumption is justified since the concentration of water is essentially constant. However, a word of caution is needed here as the standard Gibbs energy and enthalpy of formation for water must be explicitly included if it is a reactant. Again, the choice of standard states can readily be incorporated into the standard Gibbs energy change and standard Gibbs energy of formation. The best way to avoid ambiguity is to explicitly write the chemical reaction and the conditions (pH, salt, etc.) to which a given Gibbs energy change refers.

As a specific example, let us again consider the reaction catalyzed by hexokinase, the phosphorylation of glucose by ATP. We have previously calculated the standard

enthalpy change for this reaction from the known enthalpies of hydrolysis. This can also be done for the standard Gibbs energy change of the hexokinase reaction because the standard Gibbs energies of hydrolysis are known. As before, the “standard states” are 1 M at 1 atm, 298 K, pH 7.0, pMg 3.0, and an ionic strength of 0.25 M.

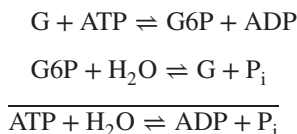
	ΔG° (kJ/mol)	ΔH° (kJ/mol)
$G + P_i \rightleftharpoons G6P + H_2O$	11.6	0.5
$ATP + H_2O \rightleftharpoons ADP + P_i$	-32.5	-30.9
$\overline{G + ATP \rightleftharpoons G6P + ADP}$	$\overline{-20.9}$	$\overline{-30.4}$

The equilibrium constant for the hexokinase reaction can be calculated from the standard Gibbs energy change:

$$\Delta G^\circ = -RT \ln K \quad \text{or} \quad K = 4630 \left\{ = \frac{[G6P][ADP]}{[G][ATP]} \right\}$$

Since the standard enthalpy change is known, -30.4 kJ/mol, the equilibrium constant can be calculated at other temperatures by use of Eq. (2-45). For example, at 310 K (37°C), the equilibrium constant is 2880. The standard entropy change for this reaction can be calculated from the relationship $\Delta G^\circ = \Delta H^\circ - T \Delta S^\circ$.

How are standard Gibbs energy changes for reactions measured? The obvious answer is by measurement of the equilibrium constant and use of Eq. (2-36). However, in order to determine an equilibrium constant, a measurable amount of both reactants and products must be present. For reactions that essentially go to completion, a sequence of reactions must be found with measurable equilibrium constants that can be summed to total the reaction of interest. For example, the equilibrium constant for the hydrolysis of ATP to ADP and P_i calculated from the aforementioned standard Gibbs energy change is 5×10^5 M under “standard conditions.” Obviously, this equilibrium constant would be very difficult to measure directly. The following reactions, however, could be used to calculate this equilibrium constant (and the standard Gibbs energy change):

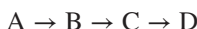


The equilibrium constant for the first reaction is about 4600 and for the second reaction is about 110 M, both of which can be measured experimentally. Tables of standard Gibbs energies of formation (Appendix 4) and standard Gibbs energy changes for reactions (Appendix 5) can be constructed using this methodology.

Knowledge of the equilibrium constants of biological reactions and their temperature dependencies is of great importance for understanding metabolic regulation. It can also be of practical importance in the design of laboratory experiments, for example, in the development of coupled assays.

3.2 METABOLIC CYCLES

Thermodynamics is particularly useful for understanding metabolism. Metabolism consists of many different sets of reactions, each set, or metabolic cycle, designed to utilize and produce very specific molecules. The reactions within a cycle are coupled in that the product of one reaction becomes the reactant for the next reaction in the cycle. These coupled reactions can be very conveniently characterized using thermodynamic concepts. Before considering a specific metabolic cycle, let us consider some general thermodynamic properties of coupled reactions. As a very simple illustration, consider the coupled reactions



The Gibbs energy changes for the first three reactions can be written as

$$A \rightarrow B \quad \Delta G_{AB} = \Delta G_{AB}^{\circ} + RT \ln \left(\frac{[B]}{[A]} \right)$$

$$B \rightarrow C \quad \Delta G_{BC} = \Delta G_{BC}^{\circ} + RT \ln \left(\frac{[C]}{[B]} \right)$$

$$C \rightarrow D \quad \Delta G_{CD} = \Delta G_{CD}^{\circ} + RT \ln \left(\frac{[D]}{[C]} \right)$$

The Gibbs energy for the overall conversion of A to D can be obtained by adding these Gibbs energies:

$$A \rightarrow D \quad \Delta G_{AD} = \Delta G_{AB}^{\circ} + \Delta G_{BC}^{\circ} + \Delta G_{CD}^{\circ} + RT \ln \left\{ \frac{([B][C][D])}{([A][B][C])} \right\}$$

$$\Delta G_{AD} = \Delta G_{AD}^{\circ} + RT \ln \left(\frac{[D]}{[A]} \right)$$

Whether A can be converted to D depends on the standard Gibbs energies for the three individual reactions and the concentrations of A and D. The concentrations of B and C are of no consequence! Note that since the total standard Gibbs energy determines whether A will be converted to D, it is possible for one of the standard Gibbs energy changes of the intermediate steps to be very unfavorable (positive) if it is balanced by a very favorable (negative) standard Gibbs energy change.

Metabolic pathways contain hundreds of different reactions, each catalyzed by a specific enzyme. Although the thermodynamic analysis shown earlier indicates that only the initial and final states need be considered, it is useful to analyze a metabolic pathway to see how the individual steps are coupled to each other through the associated Gibbs energy changes. The specific metabolic pathway that we will examine is anaerobic glycolysis. Anaerobic glycolysis is the sequence of reactions that metabolizes glucose into lactate and also produces ATP, the physiological energy currency. As we have seen, the standard Gibbs energy for the hydrolysis of ATP is

TABLE 3-1. Gibbs Energy Changes for Anaerobic Glycolysis

Reaction	ΔG° (kJ/mol)	ΔG (kJ/mol)
<i>Part One</i>		
Glucose + ATP $\rightleftharpoons$ Glucose-6-P + ADP	-16.7	-33.3
Glucose-6-P $\rightleftharpoons$ Fructose-6-P	1.7	-2.7
Fructose-6-P + ATP $\rightleftharpoons$ Fructose-1,6-bisphosphate + ADP	-14.2	-18.6
Fructose-1-6,-bisphosphate $\rightleftharpoons$ Dihydroxyacetone-P + Glyceraldehyde- 3 -P	23.9	0.7
Dihydroxyacetone-P $\rightleftharpoons$ Glyceraldehyde-3-P	7.5	2.6
Glucose + 2 ATP $\rightleftharpoons$ 2 ADP + 2 Glyceraldehyde-3-P	2.2	-51.3
<i>Part Two</i>		
Glyceraldehyde-3-P + P _i + NAD ⁺ $\rightleftharpoons$ 1,3-Bisphosphoglycerate + NADH	6.3	-1.0
1,3-Bisphosphoglycerate + ADP $\rightleftharpoons$ 3-P-Glycerate + ATP	-18.9	-0.6
3-Phosphoglycerate $\rightleftharpoons$ 2-Phosphoglycerate	4.4	1.0
2-Phosphoglycerate $\rightleftharpoons$ Phosphoenolpyruvate + H ₂ O	1.8	1.1
Phosphoenolpyruvate + ADP $\rightleftharpoons$ Pyruvate + ATP	-31.7	-23.3
Pyruvate + NADH $\rightleftharpoons$ Lactate + NAD ⁺	-25.2	1.9
Glyceraldehyde-3P + P _i + 2 ADP $\rightleftharpoons$ Lactate + 2 ATP + H ₂ O	-63.3	-20.9

Source: Adapted from R. H. Garrett and C. M. Grisham, *Biochemistry*, Saunders College Publishing, Philadelphia, 1995, pp. 569–597.

quite large and negative so that the hydrolysis of ATP can be coupled to reactions with an unfavorable Gibbs energy change. The sequence of reactions involved in anaerobic glycolysis is shown in Table 3-1, along with the standard Gibbs energy changes. The standard states are as usual for biochemical reactions, namely, pH 7 and 1 M for reactants, with the concentration of each reactant as the sum of all ionized species. The activity of water is assumed to be unity. (This is somewhat different from Appendices 4 and 5, where ionic strength and, in some cases, the magnesium ion concentration are specified.)

The overall reaction for anaerobic glycolysis is



In anaerobic metabolism, the pyruvate produced in the second to the last step is converted to lactate in muscle during active exercise. In aerobic metabolism, the pyruvate that is produced in the second to the last step is transported to the mitochondria, where it is oxidized to carbon dioxide and water in the citric acid cycle. The reactions in Table 3-1 can be divided into two parts. The first part produces 2 moles of glyceraldehyde-3-phosphate according to the overall reaction

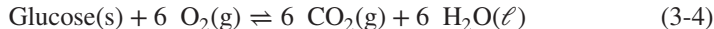


Note that this reaction actually requires 2 moles of ATP. However, the second part of the cycle produces 4 moles of ATP with the overall reaction



Since 2 moles of glyceraldehyde-3-phosphate are produced in part one of the cycle, Eq. (3-3) must be multiplied by 2 and added to Eq. (3-2) to give the overall reaction, Eq. (3-1). The standard Gibbs energies, of course, also must be multiplied by 2 for the reactions in part two of glycolysis and added to those for the reactions in part one to give the overall standard Gibbs energy change, -124.4 kJ/mol . If the standard Gibbs energies of formation in Appendix 4 are used to calculate the standard Gibbs energy change for Eq. (3-3), a value of -128.6 kJ/mol is obtained. This small difference can be attributed to somewhat different standard states. Unfortunately, not all of the necessary Gibbs energies of formation are available in Appendix 4 to calculate standard Gibbs energy changes for all of the reactions in Table 3-1.

The standard enthalpy change for Eq. (3-3) is -63.0 kJ/mol , so that a substantial amount of heat is produced by glycolysis. The standard entropy change can be calculated from the known standard Gibbs energy and enthalpy changes and is 220 J/(mol K) . Thus, both the standard enthalpy and entropy changes are favorable for the reaction to proceed. It is interesting to compare these numbers with those for the direct oxidation of glucose:



The standard Gibbs energy change for this reaction is -2878.4 kJ/mol , the standard enthalpy change is -2801.6 kJ/mol , and the standard entropy change is -259 J/(mol K) . This process produces a very large amount of heat relative to that produced by the metabolic cycle. This would not be very useful for physiological systems.

As we have stressed previously, for a single reaction, it is not the standard Gibbs energy change that must be considered in determining whether products are formed, it is the Gibbs energy for the particular concentrations of reactants that are present. In order to calculate the Gibbs energy changes for the reactions in Table 3-1, the concentrations of metabolites must be known. These concentrations have been determined in erythrocytes and are summarized in Table 3-2. The Gibbs energy changes calculated with these concentrations using Eq. (2-35) and the standard Gibbs energy changes are included in Table 3-1. The additional assumptions have been made that these concentrations are valid at 298 K , although they were determined at 310 K , and $[\text{NADH}]/[\text{NAD}^+] = 1.0 \times 10^{-3}$. (We have elected to carry out calculations at 298 K , where the standard Gibbs energies are known, rather than at the physiological temperature of 310 K . This does not alter any of the conclusions reached.)

Consideration of Gibbs energy changes, rather than standard Gibbs energy changes, produces some interesting changes. The overall standard Gibbs energy change for the first part of glycolysis is $+2.2 \text{ kJ/mol}$, whereas the Gibbs energy change is -51.3 kJ/mol . By contrast, the standard Gibbs energy change for

TABLE 3-2. Steady-State Concentrations for Glycolytic Intermediates in Erythrocytes

Metabolite	Concentration (mM)
Glucose	5.0
Glucose-6-phosphate	0.083
Fructose-6-phosphate	0.014
Fructose-1,6-bisphosphate	0.031
Dihydroxyacetone phosphate	0.14
Glyceraldehyde-3-phosphate	0.019
1,3-Bisphosphoglycerate	0.001
2,3-Bisphosphoglycerate	4.0
3-Phosphoglycerate	0.12
2-Phosphoglycerate	0.030
Phosphoenolpyruvate	0.023
Pyruvate	0.051
Lactate	2.9
ATP	1.85
ADP	0.14
P _i	1.0

Source: Adapted from S. Minakami and H. Yoshikawa, *Biochem. Biophys. Res. Commun.* **18**, 345 (1965).

the second part of glycolysis is much more negative than the Gibbs energy change. Thus, when considering the coupling of chemical reactions, considerable care must be exercised in making comparisons. Of course, as we stated at the beginning, the concentrations of the intermediates are of no consequence in determining the overall Gibbs energy changes. The reader might wish to confirm that this is indeed the case.

Before we leave our discussion of glycolysis, it is worth addressing the individual reactions in the cycle. All of the reactions are catalyzed by enzymes. If this were not the case, the reactions would occur much too slowly to be physiologically relevant.

The first step is the very favorable phosphorylation of glucose—both the standard Gibbs energy change and the Gibbs energy change are favorable. The advantage to the cell of phosphorylating glucose is that creating a charged molecule prevents it from diffusing out of the cell. Furthermore, the intracellular concentration of glucose is lowered so that if the concentration of glucose is high on the outside of the cell, more glucose will diffuse into the cell. The second step, the isomerization of glucose-6-phosphate to fructose-6-phosphate, has a somewhat unfavorable standard Gibbs energy change, but the Gibbs energy change is favorable enough for the reaction to proceed. The next step has a very favorable standard Gibbs energy change, as well as a favorable Gibbs energy change, because the phosphorylation of fructose-6-phosphate is coupled to the hydrolysis of ATP. This very irreversible step is the commitment by the cell to metabolize glucose, rather

than to store it. The next two steps produce glyceraldehyde-3-phosphate, the fuel for the second half of glycolysis that produces 4 moles of ATP. Both of these reactions have very unfavorable standard Gibbs energy changes, although the Gibbs energy changes are only slightly positive. This completes the first part of glycolysis. Note that 2 moles of ATP have been utilized to produce the final product, glyceraldehyde-3-phosphate. As noted previously, the standard Gibbs energy change for this first part is actually unfavorable, but the Gibbs energy change is quite favorable.

The purpose of the second part of glycolysis is to convert a substantial portion of the metabolic energy of glucose into ATP. In comparing the standard Gibbs energy changes with the Gibbs energy changes of the individual steps, it is worth noting that the second step, the formation of 3-phosphoglycerate, has a very favorable standard Gibbs energy change, yet the Gibbs energy change is approximately zero so that this reaction is approximately at equilibrium. This is true of essentially all of the reactions in this part of glycolysis, except for the reaction that produces pyruvate, where both the standard Gibbs energy change and the Gibbs energy change are quite negative. Both the overall standard Gibbs energy change and the Gibbs energy change for the second part of glycolysis are very favorable.

This thermodynamic analysis of glycolysis is a good example of how thermodynamics can provide a framework for understanding the many coupled reactions occurring in biology. It also illustrates how metabolism is utilized to produce molecules such as ATP that can be used to drive other physiological reactions, rather than converting most of the Gibbs energy to heat.

3.3 DIRECT SYNTHESIS OF ATP

As we have seen, the standard Gibbs energy change for the hydrolysis of ATP to ADP and P_i is -32.5 kJ/mol, so it seems unlikely that ATP would be synthesized by the reverse of this reaction. The concentrations of reactants cannot be adjusted sufficiently to make the overall Gibbs energy favorable. However, we know that ATP is synthesized directly from ADP and P_i in mitochondria. For many years, people in this field grappled with how this might happen: Both probable and improbable mechanisms were proposed. In 1961, Peter Mitchell proposed that the synthesis of ATP occurred due to a coupling of the chemical reaction with a proton gradient across the membrane (1). This hypothesis was quickly verified by experiments and he received a Nobel Prize for this work.

The enzyme responsible for ATP synthesis, ATP synthase, consists of a protein “ball,” which carries out the catalytic function, coupled to membrane-bound proteins through which protons can be transported. The process of ATP synthesis is shown schematically in Figure 3-1. Although this enzyme is found only in mitochondria in humans, it is quite ubiquitous in nature and is found in chloroplasts, bacteria, and yeast, among others. The chemiosmotic hypothesis states that a pH gradient is established across the membrane by a series of electron transfer reactions and that ATP

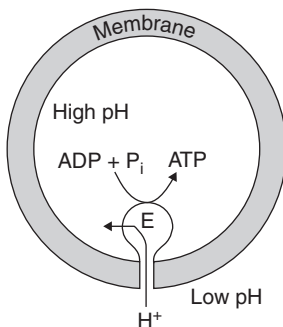
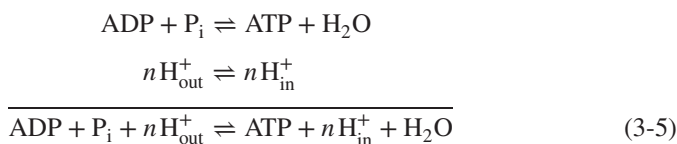


FIGURE 3-1. Schematic representation of ATP synthase, E, in mitochondria. The enzyme structure inside the mitochondria contains the catalytic sites. Protons are pumped from the outside of the membrane to the inside as ATP is synthesized.

synthesis is accompanied by the simultaneous transport of protons across the membrane. The overall reaction can be written as the sum of two reactions:



The value of n has been determined to be 3 (see Ref. 2). The Gibbs energy change for the transport of protons in Eq. (3-5) can be written as

$$\Delta G = 3RT \ln \left(\frac{[\text{H}_{\text{in}}^+]}{[\text{H}_{\text{out}}^+]}\right) \quad (3-6)$$

(The standard Gibbs energy change for this process is zero since the standard state for the hydrogen ion is the same on both sides of the membrane.)

At 298 K, a pH differential of one unit gives a Gibbs energy change of -17.1 kJ/mol. The actual physiological situation is even more favorable as a membrane potential exists whereby the membrane is more negative on the inside relative to the outside. If we utilize the extended chemical potential, Eq. (2-49), an additional term is added to Eq. (3-6) equal to $3F\Psi$, where 3 is the number of protons transported, F is the Faraday, and Ψ is the membrane potential. For a membrane potential of -100 mV, -29 kJ would be added to the Gibbs energy change in Eq. (3-6).

The standard Gibbs energy for the synthesis of ATP from ADP and P_i is $+32.5$ kJ/mol, but we need to know the Gibbs energy change under physiological conditions. Although the concentrations of the reactants are not known exactly, we can estimate that the ratio of ATP to ADP is about 100, and the concentration of phosphate is $1 - 10$ mM. This makes the ratio $[\text{ATP}]/([\text{ADP}][\text{P}_i])$ equal to 100–1000.

The Gibbs energy change at 298 K is

$$\Delta G = 32.5 + RT \ln(100 - 1000) = 32.5 + (11.4 - 17.1) = 43.9 - 49.6 \text{ kJ/mol}$$

Thus, the coupling of the synthesis of ATP to a modest proton gradient and membrane potential can readily provide the necessary Gibbs energy for the overall reaction to occur.

The principle to be learned from this example is that the coupling of Gibbs energies is very general. It can involve chemical reactions only, as in glycolysis, or it can involve other processes such as ion transport across membranes, as in this example.

3.4 ESTABLISHMENT OF MEMBRANE ION GRADIENTS BY CHEMICAL REACTIONS

In discussing ATP synthesis, we have not specified how the electrochemical gradient is established. This involves a complex sequence of coupled reactions that we shall not discuss here. Instead, we will discuss a case where the Gibbs energy associated with the hydrolysis of ATP is used to establish an ion gradient. The process of signal transduction in the nervous system involves the transport of Na^+ and K^+ across the membrane. Neuronal cells accumulate K^+ and have a deficit of Na^+ relative to the external environment. (This is also true for other mammalian cells.) When an electrical signal is transmitted, this imbalance is altered. The imbalance causes a resting membrane potential of about -70 mV . This situation is illustrated in Figure 3-2, with some typical ion concentrations.

How is the ion gradient established? This is done by a specific ATPase, the Na^+/K^+ ATPase, that simultaneously pumps ions and hydrolyzes ATP. The process can be

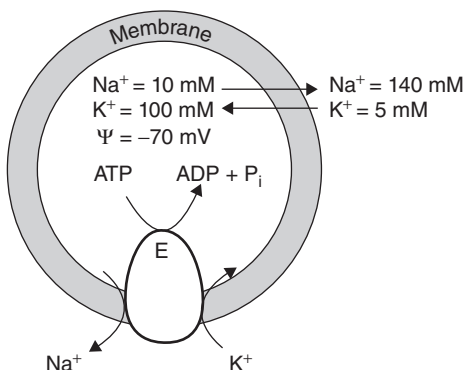
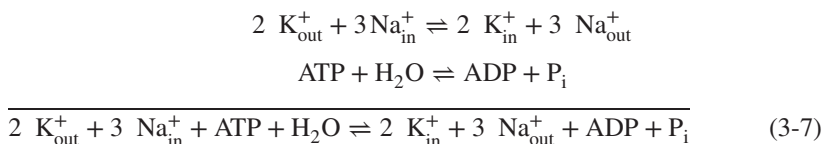


FIGURE 3-2. Schematic representation of the Na^+ and K^+ gradients in a cell. The Gibbs energy necessary for the creation of these gradients is generated by the enzymatic hydrolysis of ATP. The Na^+/K^+ ATPase is designated as E, and typical concentrations of the ions and membrane potential, Ψ , are given.

written as the sum of two reactions:



The stoichiometry has been established by experimental measurements. The Gibbs energy change for the first step is

$$\Delta G = RT \ln \left\{ \frac{([\text{Na}_{\text{out}}^{+}]^3 [\text{K}_{\text{in}}^{+}]^2)}{([\text{Na}_{\text{in}}^{+}]^3 [\text{K}_{\text{out}}^{+}]^2)} \right\} + 3F\Psi - 2F\Psi$$

With $\Psi = 70 \text{ mV}$, $T = 298 \text{ K}$, and the concentrations of Na^{+} and K^{+} in Figure 3-2, $\Delta G = 41.3 \text{ kJ/mol}$. Note that the membrane potential produces a favorable (negative) Gibbs energy change for the transport of K^{+} and an unfavorable Gibbs energy change (positive) for the transport of Na^{+} . We have previously calculated that ΔG for ATP hydrolysis is $-(43.9-49.6) \text{ kJ/mol}$. Therefore, the hydrolysis of ATP is sufficient to establish the ion gradient and accompanying membrane potential. This process is called *active transport*.

The coupling of Gibbs energies in biological systems can be used to understand many of the events occurring, and many other interesting examples exist. However, we will now turn to consideration of protein and nucleic acid structures in terms of thermodynamics.

3.5 PROTEIN STRUCTURE

An extensive discussion of protein structure is beyond the scope of this treatise, but various aspects of protein structure are considered throughout this text. Many excellent discussions of protein structure are available (cf. Refs. 3-5). A few aspects of protein structure are now discussed in terms of thermodynamic considerations. Proteins, of course, are polymers of amino acids and the amino acids contain both polar and nonpolar groups. The structures of the 20 common amino acids are given in Appendix 2, and the covalent structure of a polypeptide chain is shown in Figure 3-3. We first consider the role of nonpolar groups in proteins, that is, amino acid side chains containing methylene and methyl groups and aromatic residues.

As a starting point, let us consider the thermodynamics of transferring the hydrocarbons methane and ethane from an organic solvent to water. The thermodynamic parameters can be measured for this process simply by determining the solubility of the hydrocarbons in each of the solvents. Since the pure hydrocarbon is in equilibrium with each of the saturated solutions, the two saturated solutions must be in equilibrium with each other. The transfer reaction can be written as the sum of the solubility

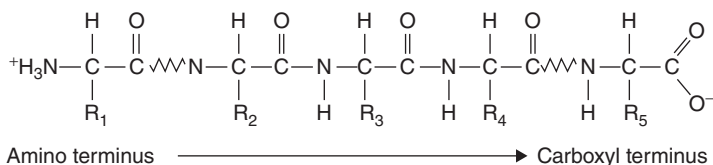
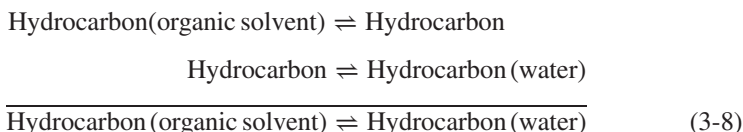


FIGURE 3-3. A polypeptide chain. The amino acid side chains are represented by R_i , and the amino and carboxyl termini are shown.

equilibria:



The transfer Gibbs energy is

$$\Delta G_t = RT \ln \left(\frac{[X_O]}{[X_W]} \right)$$

where X_O and X_W are the mole fraction solubilities in the organic solvent and water, respectively. (The use of mole fractions is the appropriate concentration scale in this instance because of standard state considerations that we will not deal with here.) Measurement of the temperature dependence of the transfer Gibbs energy will permit ΔH and ΔS to be determined (Eqs. (2-42 and 2-43)). The thermodynamic parameters obtained at 298 K in several organic solvents are summarized in Table 3-3.

As expected, the Gibbs energy change is unfavorable—hydrocarbons do not dissolve readily in water. However, the energy change is favorable so that the negative entropy change is responsible for the unfavorable Gibbs energy. What is the reason for the negative entropy change? It is due to the fact that the normal water structure is broken by the insertion of the hydrocarbon and the water molecules tend to form

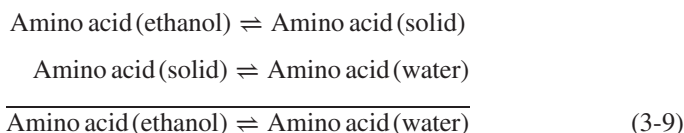
TABLE 3-3. Thermodynamic Parameters for Hydrocarbon Transfer from Organic Solvents to Water at 298 K

Transfer reaction	ΔS_t [cal/(mol K)]	ΔH_t (kcal/mol)	ΔG_t (kcal/mol)
CH ₄ , benzene to water	-18	-2.8	+2.6
CH ₄ , ether to water	-19	-2.4	+3.3
CH ₄ , CCl ₄ to water	-18	-2.5	+2.9
C ₂ H ₆ , benzene to water	-20	-2.2	+3.8
C ₂ H ₆ , CCl ₄ to water	-18	-1.7	+3.7

Source: Adapted from W. Kauzmann, *Adv. Protein Chem.* **14**, 1 (1959).

a hydrogen-bonded structure of water. Thus, it is the formation of ice-like ordered structures around the hydrocarbon that is the source of the negative entropy change.

To make this more relevant to proteins, the Gibbs energy of transfer for amino acids from a hydrocarbon-like environment (ethanol) to water has been measured. This is done by determining the solubility of the amino acids in ethanol and water. Similar to before, in both cases the amino acid in solution is in equilibrium with the solid, so that the transfer Gibbs energy can be measured:



Some of the results obtained are summarized in Table 3-4. In order to interpret these results, we must remember that in ethanol the amino acid is uncharged ($\text{NH}_2\text{—RCH—COOH}$), whereas in water it is a zwitterion ($\text{NH}_3^+ \text{—RCH—COO}^-$).

If glycine, which has no side chains, is taken as the standard, then subtracting its transfer Gibbs energy from the transfer Gibbs energies of the other amino acids will give the transfer Gibbs energy for the amino acid side chains. The overall transfer Gibbs energy for all of the amino acids is negative (favorable) because the charged amino and carboxyl groups are solvated by water, a highly favorable interaction. However, the standard Gibbs energy changes of transfer for the hydrophobic side chains are all positive, as was seen for the transfer of methane and ethane from an organic solvent into water.

What is the relevance of such data for protein structure? These data indicate that the apolar side chains of amino acid side chains would prefer to be in a nonaqueous environment. That is, they prefer to cluster together. This is, in fact, true for most proteins. The hydrophobic groups aggregate together on the interior of the protein, forming a hydrophobic core, with the more polar groups tending to be on the outside interacting with water. This important concept was enunciated by Walter Kauzmann in 1959 (6). Although the interactions associated with forming the hydrophobic core are often called hydrophobic bonds, it is important to remember that the driving force

TABLE 3-4. Gibbs Energy Changes for Transferring Amino Acids from Ethanol to Water at 298 K

Compound	ΔG_t (kcal/mol)	$\Delta G_{t,\text{side chain}}$ (kcal/mol)
Glycine	−4.63	—
Alanine	−3.90	+0.73
Valine	−2.94	+1.69
Leucine	−2.21	+2.42
Isoleucine	−1.69	+2.97
Phenylalanine	−1.98	+2.60
Tyrosine	−1.78	+2.85

Source: Adapted from C. Tanford, *J. Am. Chem. Soc.* **84**, 4240 (1962).

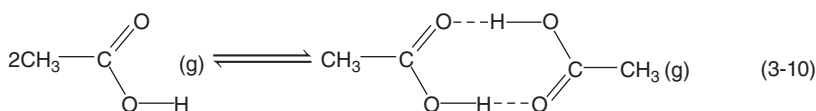
TABLE 3-5. Thermodynamic Parameters for Hydrogen Bond Formation^a

Reactant	Solvent	ΔG_{298}° (kcal/mol)	ΔH_{298}° (kcal/mol)	ΔS_{298}° [cal/(mol K)]
CH_3COOH^b	Gas	-8.02	-15.9	-26.6
$\text{CH}_3\text{CONH}_2^c$	CCl_4	-0.92	-4.2	-11
	Dioxane	0.39	-0.8	-4
	H_2O	3.1	0.0	-10

^aStandard state is 1 M.^bSee J. O. Halford, *J. Chem. Phys.* **9**, 859 (1941).^cSee I. M. Klotz and J. S. Franzen, *J. Am. Chem. Soc.* **84**, 3461 (1962).

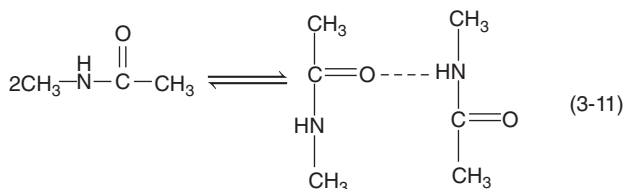
for forming a hydrophobic core is not the direct interactions between hydrophobic groups; instead, it is the release of water molecules from the ice-like structures that surround hydrophobic groups in water. The thermodynamic analysis clearly indicates that the formation of “hydrophobic bonds” is an entropy-driven process. In fact, ΔH_t for the transfer of methane and ethane from water to organic solvent is positive, that is, energetically unfavorable. This means that an increase in temperature will tend to strengthen the hydrophobic bonding—if ΔH remains positive over the temperature range under consideration.

The formation of hydrogen bonds in protein structures is prevalent, and we will now consider the thermodynamics of hydrogen bond formation. Thermodynamic studies have been made of many different types of hydrogen bonds. A few selected examples are summarized in Table 3-5. The first is the formation of a hydrogen-bonded dimer of acetic acid molecules in the gas phase:



As expected, both ΔH° and ΔS° are negative. The former represents the energy produced in forming two hydrogen bonds and the latter is due to two molecules forming a dimer (a more ordered system). As two hydrogen bonds are formed, the enthalpy change associated with a single hydrogen bond is about -7 kcal/mol.

The second example is the dimerization of *N*-methylacetamide, a good model for hydrogen bonding involving the peptide bond:



This dimerization was studied in a variety of solvents, and the results in carbon tetrachloride, dioxane, and water are included in Table 3-5. In carbon tetrachloride, the solvent does not compete for the hydrogen bonds of *N*-methylacetamide, and the results are not very different from those for acetic acid dimerization in the gas phase, namely, ΔH° for formation of a single hydrogen bond is about -4 kcal/mol and ΔS° is about -11 cal/(mol K). However, in water, which can also form hydrogen bonds with *N*-methylacetamide and is present at a concentration of about 55 M, ΔH° is approximately zero. This indicates that there is not a significant enthalpy difference between the water-*N*-methylacetamide hydrogen bond and the *N*-methylacetamide-*N*-methylacetamide hydrogen bond. Note that the standard Gibbs energy change in water is 3 kcal/mol, so that the amount of dimer formed is very small even at 1 M *N*-methylacetamide. Dioxane has two oxygens that can accept a hydrogen bond so that ΔH° is only slightly negative.

The conclusion reached from these data (and considerably more data not presented) is that the stability of a water-protein hydrogen bond is similar to that of intramolecular protein hydrogen bonds. Therefore, a single intramolecular hydrogen bond on the surface of the protein is unlikely to be a strong stabilizing factor. On the other hand, if the hydrophobic interior of a protein excludes water, a strong hydrogen bond might exist in the interior of a protein. This statement is complicated by the fact that proteins have "breathing" motions; that is, the protein structure continually opens and closes with very small motions so that completely excluding water from the interior may be difficult.

The role of the hydrogen bond in stabilizing proteins is still a matter of some debate. The statement is often made that hydrophobic interactions are the primary source of protein stability and hydrogen bonding provides specificity (6). There is no doubt that extended hydrogen-bonded systems are extremely important structural elements in proteins. Two examples are given in Figures 3-4 and 3-5, the α -helix and the β -pleated sheet. The α -helix is a spiral structure with 3.6 amino acids per turn of the helix. Every peptide carbonyl is a hydrogen bond acceptor for the peptide N-H four residues away. This structure is found in many different proteins. The β -pleated sheet is also a prevalent structure in proteins. In the β -pleated sheet, each chain is a "pleated" helix. All of the peptide bonds participate in hydrogen bonding, but the bonds are all between chains, rather than intrachain.

Finally, we will say a few words about the role of electrostatic interactions in protein structures. The discussion will be confined to charge-charge interactions, even though more subtle interactions involving, for example, dipoles and/or induced dipoles are important. Many of the amino acids have side chains with ionizable groups, so that a protein contains many acids and bases. For example, a carboxyl group can ionize according to the scheme



Here P designates the protein. Although not shown, water plays an extremely important role in this equilibrium. Water is a strong dipole and strongly solvates ions, forming a hydration "sheath." Acetic acid is a reasonable model for

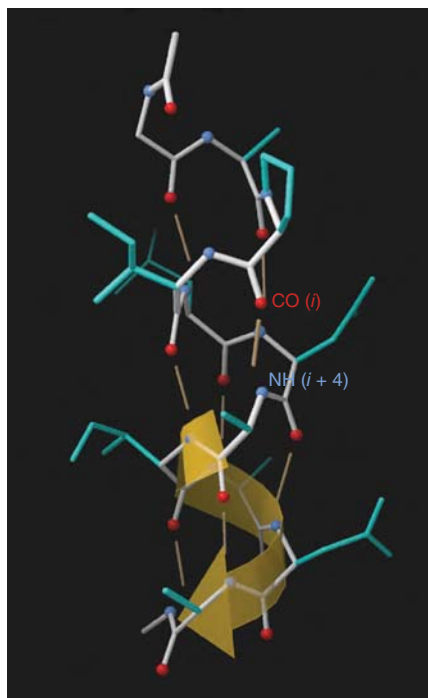


FIGURE 3-4. The α -helix found in many proteins. The yellow arrow follows the right-handed spiral of one helical turn. The hydrogen bonds between backbone peptide bonds are brown lines, the oxygens are red, and the nitrogens blue. The hydrogen bonds are formed between the i th carbonyl and $i + 4$ NH in the peptide backbone. Copyright by Professor Jane Richardson. Reprinted with permission.

this reaction: The thermodynamic parameters characterizing its ionization are $\Delta G^\circ = 6.6 \text{ kcal/mol}$, $\Delta H^\circ = 0$, and $\Delta S^\circ = -22 \text{ cal/(mol K)}$. The negative entropy change is due to the ordering of water molecules around the ions. Note that the ionization process is thermally neutral. The enthalpy changes associated with the solvation of ions are generally negative but, in this case, are balanced by the enthalpic change associated with breaking the oxygen-hydrogen bond.

The ionization constants for ionizable groups on proteins are generally not the same as those of simple model compounds. This is because the ionization process is influenced by the charge of the protein created by other ionizable groups and, in some cases, by special structural features of the protein. This factor can be included explicitly in a thermodynamic analysis of protein ionizations by writing the Gibbs energy associated with ionization as the sum of the Gibbs energy for the model compound, or the intrinsic Gibbs energy, and the Gibbs energy of interaction:

$$\Delta G_{\text{ionization}} = \Delta G_{\text{intrinsic}}^\circ + \Delta G_{\text{interaction}} \quad (3-13)$$

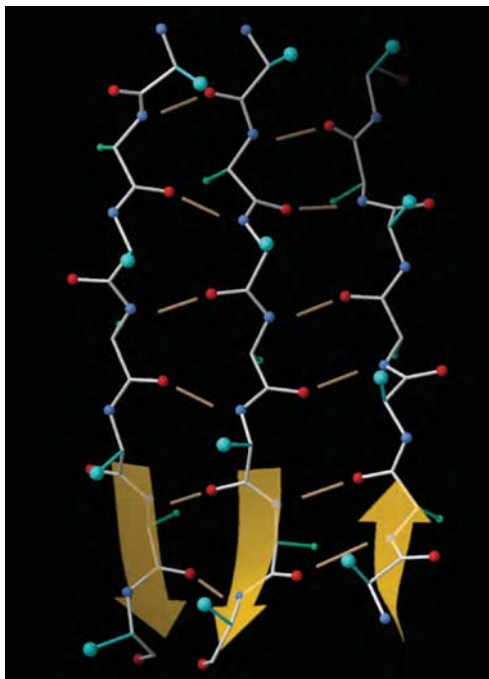


FIGURE 3-5. The β -pleated sheets found in many proteins. Both parallel and antiparallel strand-strand interactions are shown, as indicated by the yellow arrows. Again, the hydrogen bonds between backbone peptide bonds are shown as brown lines between the oxygens and nitrogens. Copyright by Professor Jane Richardson. Reprinted with permission.

The ionization properties of proteins have been studied extensively, both experimentally and theoretically. For our purposes, it is important to recognize that the strong solvation of ions means that charged groups will tend to be on the outside of the protein, readily accessible to water.

A reasonable question to ask is: If there are so many charged groups on proteins, would they not influence the structure simply because charged groups of opposite sign attract, and those with the same sign repel? This is certainly the case—at very high pH, a protein becomes very negatively charged and the interactions between negative charges can eventually cause the native structure to disappear. Similarly, at very low pH, the interactions between positive charges can cause disruption of the native structure. Conversely, the formation of a salt linkage between groups with opposite charges can stabilize structures. We can make a very simple thermodynamic analysis of charge–charge interactions by recognizing that the Gibbs energy of formation of an ion pair is simply the work necessary to bring the ions to within a specified distance, a :

$$\Delta G = \frac{z_1 z_2 e^2}{Da} \quad (3-14)$$

Here the z_i are the ionic charge numbers, $-e$ is the charge of an electron, and D is the dielectric constant. The assumption of a simple Coulombic potential is a gross oversimplification but suffices for our purposes. A more complete potential would include the ionic environment of the medium, the structure of the ions, and the microscopic structure of the solvent. If the charges are assumed to be 1, the distance of closest approach 4 Å, and the dielectric constant of water 80, ΔG is about 1 kcal/mol. Thus, the interaction energy is not very large in water. However, the interior of a protein is more like an organic solvent and organic solvents have dielectric constants of about 3, which would significantly increase the Gibbs energy of interaction. In fact, salt linkages are rarely found near the surface of proteins but are found in the interior of proteins, as expected.

If the Gibbs energy of ionic interactions is given by Eq. (3-14), the enthalpy and entropy can easily be calculated from Eqs. (2-42) and (2-43). Since only the dielectric constant in Eq. (3-14) is temperature dependent, this gives

$$\Delta H = -T^2 \left[\frac{d\left(\frac{\Delta G}{T}\right)}{dT} \right] = \left(\frac{\Delta G}{D}\right) \left[\frac{d(DT)}{dT} \right] \quad (3-15)$$

$$\Delta S = -\frac{d\Delta G}{dT} = \left(\frac{\Delta G}{D}\right) \left(\frac{dD}{dT}\right) \quad (3-16)$$

In water, both of these derivatives are negative so that if ΔG is negative, both the entropy and enthalpy changes are positive. The positive entropy change can be rationalized as being due to the release of water of hydration of the ions when the ion pair is formed. The enthalpy change is a balance between the negative enthalpy change from bringing the charges closer and the positive enthalpy change associated with removing the hydration shell. Since the enthalpy change is positive, the strength of the ion pair interaction will increase as the temperature is increased—exactly as for hydrophobic interactions.

Although this is an abbreviated discussion, it is clear that we know a great deal about the thermodynamics of interactions that occur in proteins. Because of this, you might think that we could examine the amino acid sequence of any protein and predict its structure by looking at the possible interactions that occur and finding the structure that has the minimum Gibbs energy. This has been a long-standing goal of protein chemistry, but we are not yet able to predict protein structures. Why is this the case? The difficulty is that there are thousands of possible hydrogen bonding interactions, ionic interactions, hydrophobic interactions, and so on. As we have seen, each of the individual interactions is associated with small Gibbs energy and enthalpy changes—in some cases, we cannot even determine the sign. The sum of the positive Gibbs energies of interactions is very large, as is the sum of the negative Gibbs energies. It is the difference between the positive and negative Gibbs energies that determines the structure. So we have two problems: accurately assessing the Gibbs energies of individual interactions and then taking the difference between two large numbers to determine which potential structure has the minimum Gibbs energy. These are formidable problems, but significant progress has been made

toward achieving the ultimate goal of predicting protein structures. The possibility also exists that the structure having the lowest Gibbs energy is not the biologically relevant structure. This may be true in a few cases but is unlikely to be a problem for most proteins.

3.6 PROTEIN FOLDING

The folding of proteins into their biologically active structures has obvious physiological importance. In addition, understanding the process in molecular detail is linked directly to understanding protein structure. The study of protein folding, and the reverse process of unfolding, is a major field of research, and we will only explore a few facets of this fascinating subject. For our discussion, we will concentrate on protein unfolding as this is most easily experimentally accessible. There are many ways of unfolding proteins. When the temperature increases, proteins will eventually unfold. From a thermodynamic standpoint, this is because the $T\Delta S$ term eventually dominates in determining the Gibbs energy change, and we know that the unfolded state is more disordered than the native state at sufficiently high temperatures.

Chemical denaturants such as acid, base, urea, and guanidine chloride are also often used to unfold proteins. The role of a neutrally charged denaturant such as urea can be understood in thermodynamic terms by considering the Gibbs energy of transfer of amino acids from water to urea, again using glycine as a reference. Some representative Gibbs energies of transfer for hydrophobic side chains are given in Table 3-6. Note that the Gibbs energies are all negative, so that removing hydrophobic side chains from the interior of the protein into 8 M urea is a favorable process.

Protein unfolding can be monitored by many different methods. Probably the most common is circular dichroism in the ultraviolet, which is quite different for native and unfolded structures. Many other spectral and physical methods work equally well. (Spectroscopic methods are discussed later in this text.) A representative plot of the fraction of denatured protein, f_D , versus temperature is shown in Figure 3-6 for the N-terminal region (amino acid residues 6–85) of λ repressor, a protein from λ phage that binds to DNA and regulates transcription. Results are shown for both no urea and 2 M urea. As expected, the protein is less stable in the presence of 2 M urea. The fraction of protein denatured can be written as

$$f_D = \frac{[D]}{([D] + [N])} = \frac{K}{(1 + K)} \quad (3-17)$$

where $[D]$ is the concentration of denatured species, $[N]$ is the concentration of native species, and $K(= [D]/[N])$ is the equilibrium constant for denaturation. This assumes that the unfolding process can be characterized by only two states—native and denatured. For some proteins, intermediates are formed as the protein unfolds and a more complicated analysis must be used.

Since the equilibrium constant can be calculated at any point on the curve and its temperature dependence can be measured, the thermodynamic parameters characterizing the unfolding can be determined. The thermodynamics of unfolding can also

TABLE 3-6. Transfer Gibbs Energies for Selected Amino Acids from Water to 8 M Urea at 298 K

Amino acid	ΔG_t (kcal/mol)	$\Delta G_{t, \text{side chain}}$ (kcal/mol)
Glycine	+0.10	0.0
Alanine	+0.03	−0.07
Leucine	−0.28	−0.38
Phenylalanine	−0.60	−0.70
Tyrosine	−0.63	−0.73

Source: Adapted from P. L. Whitney and C. Tanford, *J. Biol. Chem.* **237**, 1735 (1962).

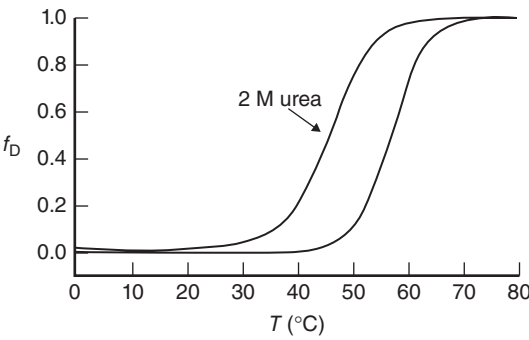


FIGURE 3-6. Fraction of denatured protein, f_D , for λ_{6-85} phage repressor protein as a function of temperature, T , in aqueous solution and in 2 M urea. Data from G. S. Huang and T. G. Oas, *Biochemistry* **35**, 6173 (1996).

conveniently be studied by scanning calorimetry. It is often found that a plot of the logarithm of the equilibrium constant versus $1/T$ is not linear, as predicted, namely,

$$\frac{d \ln K}{d(1/T)} = -\frac{\Delta H}{R} \tag{3-18}$$

This means that the enthalpy change is temperature dependent, or in other words, there is a large heat capacity difference, ΔC_p , between the native and unfolded states.

$$\frac{d\Delta H}{dT} = \Delta C_p \tag{3-19}$$

A typical experimental result for the variation of the equilibrium constant with temperature for λ repressor is given in Figure 3-7a. The plot of $\ln K$ versus $1/T$ goes through a minimum: At higher temperatures, ΔH is positive, whereas at lower temperatures, it is negative. For typical chemical reactions, this plot is a straight line! In Figure 3-7b, the standard Gibbs energy change is plotted versus the temperature. The protein is most stable at the maximum in the Gibbs energy curve, about 15°C. In Figure 3-7c, ΔH° and $T \Delta S^\circ$ are plotted versus T . Note that both ΔH° and ΔS°

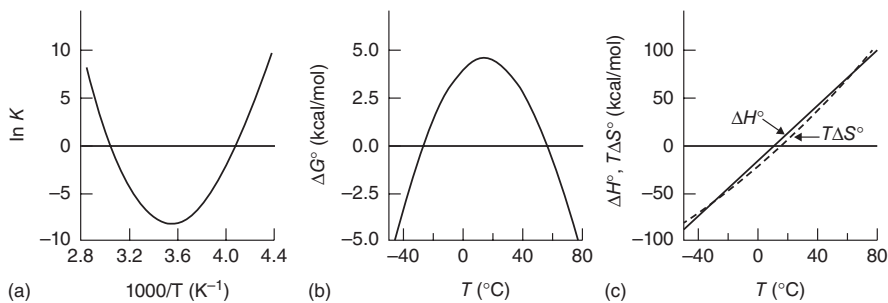


FIGURE 3-7. (a) Plot of the natural logarithm of the equilibrium constant, K , for the denaturation of λ_{6-85} phage repressor protein in aqueous solution versus the reciprocal temperature on the Kelvin scale, $1/T$. (b) Plot of ΔG° for the denaturation of λ_{6-85} phage repressor protein versus the temperature, T . (c) Plots of ΔH° and $T \Delta S^\circ$ versus the temperature, T , for the denaturation of λ_{6-85} phage repressor protein. Data from G. S. Huang and T. G. Oas, *Biochemistry* **35**, 6173 (1996).

are zero at a specific temperature. The practical implication of the positive value of ΔC_p is that the protein unfolds at both high temperatures and low temperatures. This unexpected result is not confined to λ repressor. Most proteins will unfold (denature) at low temperatures, but in many cases the temperatures where cold denaturation is predicted to occur are below 0°C .

Some thermodynamic parameters characterizing protein denaturation are given in Table 3-7 for a few proteins. In all cases, a large positive value of ΔC_p is observed. Note that at room temperature and above, ΔH° is typically large and positive, as is ΔS° , so that unfolding is an entropically driven process. Before we leave our discussion of protein folding/unfolding, one more important point should be mentioned. The transition from folded to unfolded states usually occurs over a fairly small change in temperature or denaturant concentration. This is because folding/unfolding is a highly *cooperative* process—once it starts, it proceeds with very small changes in temperature or denaturant. Cooperative processes are quite prevalent in biological systems,

TABLE 3-7. Representative Thermodynamic Parameters for Thermal Protein Denaturation in Aqueous Solution at 298 K

Protein	ΔG° (kJ/mol)	ΔH° (kJ/mol)	ΔS° [J/(mol K)]	ΔC_p [kJ/(mol K)]
Barnase	48.9	307	866	6.9
Chymotrypsin	45.7	268	746	14.1
Cytochrome <i>c</i>	37.1	89	174	6.8
Lysozyme	57.8	242	618	9.1
Ribonuclease A	27.0	294	896	5.2
λ Repressor ₆₋₈₅ ^a	17.7	90.4	244	6.0

Source: Adapted from G. I. Makhatadze and P. L. Privalov, *Adv. Protein Chem.* **47**, 307 (1995).

^aSee G. S. Huang and T. G. Oas, *Biochemistry* **35**, 6173 (1996).

particularly when regulation of the process is desired; cooperative processes will be discussed again in Chapter 17.

3.7 NUCLEIC ACID STRUCTURES

We will now briefly consider the structure of nucleic acids in terms of thermodynamics. Only a few examples will be considered. Many more complete discussions are available (3,4,7). Let us start with the well-known structure of DNA. Most DNAs consist of two chains that form a double helix. Each chain is a polymer of nucleosides linked by phosphodiester bonds as shown in Figure 3-8. Each nucleoside contains a 2'-deoxyribose sugar and a base, almost always adenine (A), thymine (T), cytosine (C), or guanine (G). The phosphodiester linkage is through the 5' and 3' positions on the sugars. Structures of these components are given in Appendix 3. By convention, a DNA chain is usually written so that the 5' end of the molecule is on the left and the 3' on the right. The B form of the DNA double helix is shown in Figure 3-9. The chains are arranged in an antiparallel fashion and form a right-handed helix. The bases are paired through hydrogen bonds on the inside of the double helix and as might be expected the negatively charged phosphodiester is on the outside. Note the similarity to protein structure as the more hydrophobic groups tend to be on the inside

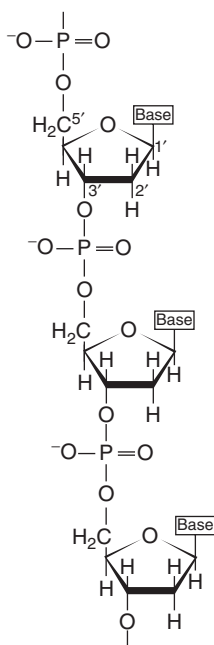


FIGURE 3-8. Structural formula of part of a DNA/RNA chain. In DNA, the bases are usually A, T, C, and G. In RNA, the T is replaced by U, and the 2'-H below the plane of the sugar is replaced by OH.

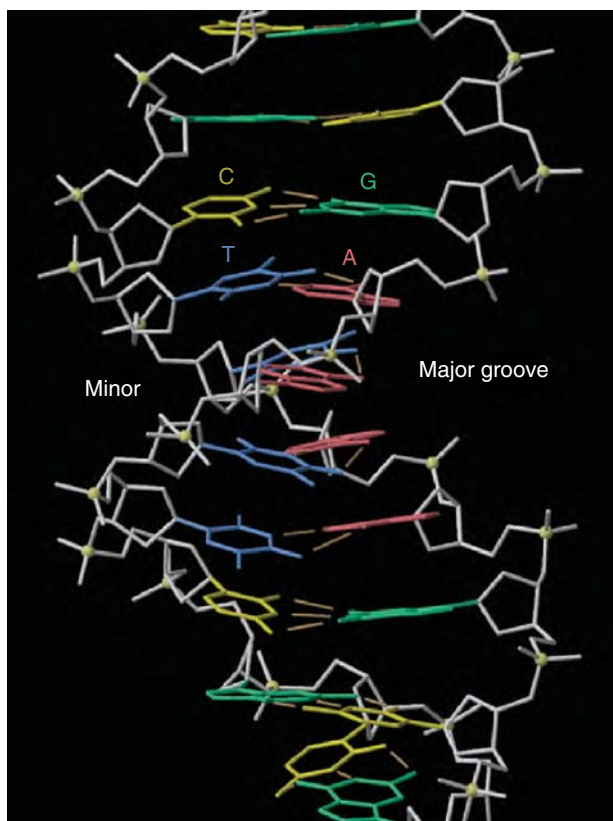


FIGURE 3-9. Stick-figure representation of the B form of the DNA double helix. The planes of the hydrogen-bonded bases can be seen, as well as the twisting of the chains to form a double helix. The pale yellow spheres are the phosphorous atoms of the sugar phosphate backbone on the outside of the helix. The bases are color coded, and the two grooves in the structure are labeled. Copyright by Professor Jane Richardson. Reprinted with permission.

and the hydrated polar groups on the outside. However, DNA is not globular, unlike most proteins, and forms a rod-like structure in isolation. These rods can bend and twist to form very compact structures in cells.

Let us first examine the hydrogen-bonded pairs in DNA. Early in the history of DNA, Erwin Chargaff noted that the fraction of bases that were A was approximately equal to the fraction that were T in many different DNAs. Similarly, the fraction of bases that were G was equal to the fraction that were C. This finding was important in the postulation of the double helix structure by James Watson and Francis Crick. As shown in Figure 3-10, these bases form hydrogen-bonded pairs, with the A–T pair forming two hydrogen bonds and the G–C pair forming three hydrogen bonds. These are not the only possible hydrogen-bonded pairs that can be formed between the four bases. Table 3-8 gives the thermodynamic parameters associated with the formation

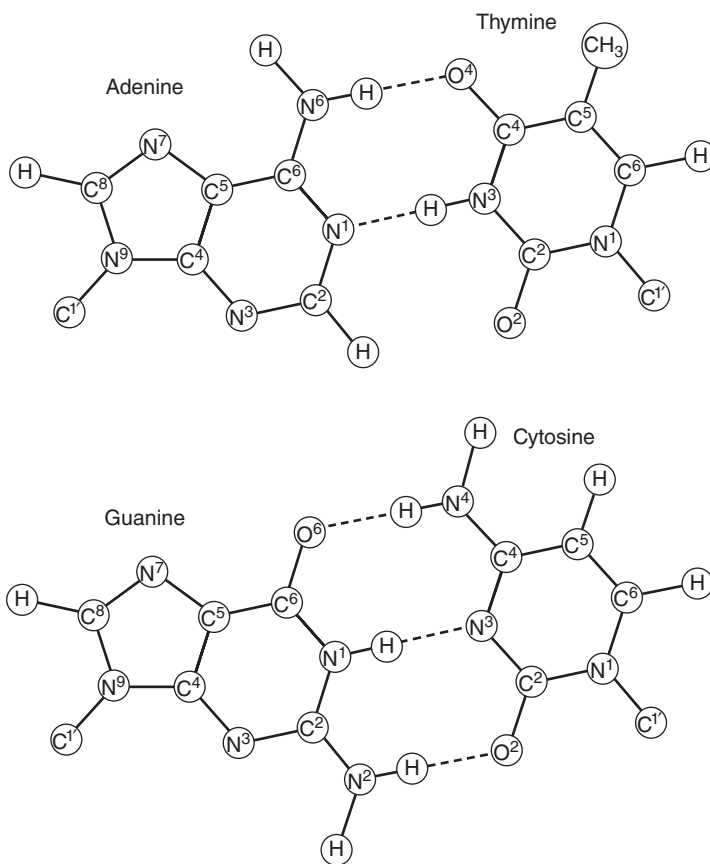


FIGURE 3-10. Formation of the hydrogen-bonded base pairs, A–T and G–C, in DNA. In RNA, U is substituted for T. All of the base rings are aromatic.

TABLE 3-8. Thermodynamic Parameters for Base Pairs in Deuteriochloroform at 298 K

Base pair	ΔG° (kcal/mol)	ΔH° (kcal/mol) ^a	ΔS° [cal/(mol K)]
A–A	–0.67	–4.0	–11.4
U–U	–1.07	–4.3	–11.0
A–U	–2.72	–6.2	–11.8
C–C	–1.97	–6.3	–15
G–G	–4.1 to –5.4	(–8.5 to –10)	(–15)
G–C	–5.4 to –6.8	(–10.0 to –11.5)	(–15)

Source: Adapted from C. R. Cantor and P. R. Schimmel, *Biophysical Chemistry*, Freeman, New York, 1980, p. 325. Data source: Y. Kyogoku, R. C. Lord, and A. Rich, *Biochim. Biophys. Acta* **179**, 10 (1969).

^aThe ΔH° values in parentheses have been calculated assuming $\Delta S^\circ = -15$ cal/(mol K).

of various hydrogen-bonded pairs in deuteriochloroform. Instead of thymine, uracil has been used—this is the base normally found in ribonucleic acids and differs only by a methyl group in the 5 position from thymine. It is clear that the A–U pair is favored over the A–A and U–U pairs and that G–C is favored over G–G and C–C although the data are not very precise. There is no simple explanation for this strong preference—it must be due to more than hydrogen bonding, perhaps an electronic effect within those specific pairs. Furthermore, the preference for A–T in the former three pairs is due to a more favorable enthalpy change. On the basis of this observation, the entropy change for the latter three pairs has been assumed to be the same in order that ΔH° can be calculated.

In water solution, the hydrogen bonding between bases is considerably weaker than in organic solvents because of the competition for hydrogen bond formation with water. The assumption is that the base pair hydrogen bonds are sufficiently shielded from water so that the structure is very stable. However, we know that this is not entirely correct as some “breathing” motions occur, so some exposure to water must also occur. Factors other than hydrogen bonding must contribute to the stability of the double helix.

A consideration comes into play in nucleic acid structures that is not a major concern in proteins, namely, the interactions between the aromatic rings of the bases. The planes of the bases lie over one another so that the π electrons interact; that is, the rings are attracted to each other. This “stacking” interaction can be studied in model systems by measuring the equilibrium constant for the interaction of nucleosides in water, where the hydrogen bonding between nucleosides is negligible. Some representative equilibrium constants for dimer formation are given in Table 3-9. The interaction is quite weak, although it appears that purine–purine interactions are stronger than purine–pyrimidine interactions, which in turn are stronger than pyrimidine–pyrimidine interactions. This is the order expected for π electron interactions. The enthalpy change associated with these interactions is somewhat uncertain but is definitely negative. A value of -3.4 kcal/mol has been obtained for formation of an A–A stack (8).

Is the “stacking” interaction entirely due to π electron interactions, or is the solvent involved, as for hydrophobic interactions in proteins? The “stacking” interaction has

TABLE 3-9. Association Constants and Standard Gibbs Energy Changes for Base Stacking in H_2O at 298 K

Base stack	K (molar $^{-1}$)	ΔG° (kcal/mol)
A–A	12	–1.50
T–T	0.91	+0.06
C–C	0.91	+0.06
T–C	0.91	+0.06
A–T	3 to 6	–0.70 to –1.10
A–C	3 to 6	–0.70 to –1.10
G–C	4 to 8	–0.80 to –1.20

Source: Adapted from T. N. Solie and J. A. Schellman, *J. Mol. Biol.* 33, 6 (1968).

been found to be solvent dependent, with water being the most favorable solvent. This suggests that hydrophobic interactions may be involved. However, you should recall that hydrophobic interactions are endothermic and entropically driven by the release of ordered water molecules around the noninteracting hydrophobic groups. The most likely possibility is that both π electron interactions and hydrophobic effects play a role in base stacking.

3.8 DNA MELTING

One of the reasons for wanting to understand the interactions in the DNA molecule is to understand the stability of DNA since it must come apart and go together as cells reproduce. One way to assess the stability of DNA is to determine its thermal stability. This can readily be studied experimentally because the ultraviolet spectra of stacked bases differ significantly from those of unstacked bases. Thus, if the temperature is raised, the spectrum of DNA changes as it “melts.” This is shown schematically in Figure 3-11. Because of the relative stability of the hydrogen bonding, the A–T

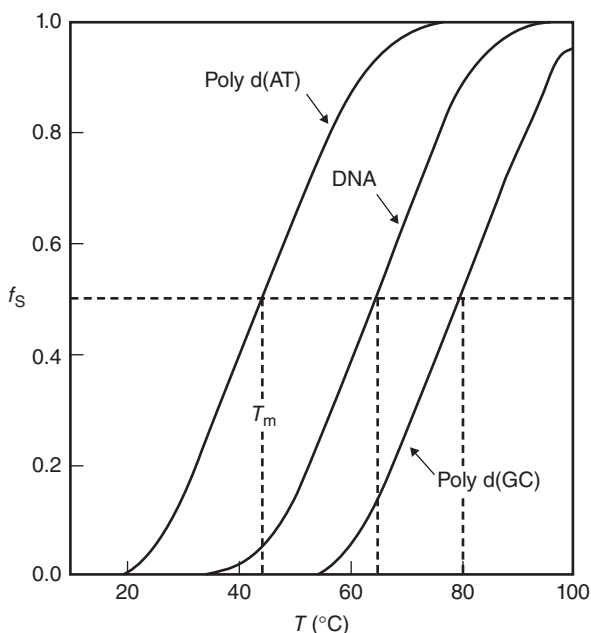


FIGURE 3-11. Hypothetical melting curves for double helix structures of poly d(AT), DNA (containing A–T and G–C base pairs), and poly d(GC). The fraction of chains not in the double helix structure, f_s , is plotted versus the temperature. The melting temperatures, T_m , are indicated by the dashed lines. The shape of the curves and the T_m values are dependent on the length of the chains and their concentrations. This drawing assumes that the concentrations and chain lengths are comparable in all three cases.

regions melt at a lower temperature than the G–C regions. The melting of AT and GC polymers can be measured, and as shown in Figure 3-11, the AT structure melts at a lower temperature than a typical DNA, and the GC structure at a higher temperature. The temperature at which half of the DNA structure has disappeared is the melting temperature, T_m . If the transition from fully helical DNA to separated fragments is assumed to be a two-state system, then thermodynamic parameters can be calculated, exactly as for protein unfolding. Thus, the equilibrium constant is 1 ($\Delta G^\circ = 0$) at T_m . Determination of the temperature dependence of the equilibrium constant permits calculation of the standard enthalpy and Gibbs energy changes for the “melting” process.

The real situation is more complex than shown, as DNA may not melt in a well-defined single step, and mixtures of homopolynucleotides can form structures more complex than dimers, but a more detailed consideration of these points is beyond the scope of this text. However, it should be mentioned that thermal melting of single-stranded homopolymers is a useful tool for studying stacking interactions: As with DNA, the spectrum will change as the bases unstack. The thermodynamic interpretation of these results is complex (cf. Ref. 3).

From a practical standpoint, many of the aforementioned difficulties can be circumvented by the use of model systems to determine the thermodynamics of adding a base pair to a chain of nucleotides. Thus, for example, an oligonucleotide A_n could be added to T_n and a complex of $A_n T_n$ formed. The same experiment can be carried out with A_{n+1} and T_{n+1} . If ΔG° is determined for the formation of each of the complexes, then the difference between the two standard Gibbs energies of formation gives the standard Gibbs energy change for the formation of one A–T pair. The temperature dependence of this Gibbs energy difference can be used to obtain the standard enthalpy and entropy changes for formation of a single A–T pair. With knowledge of the thermodynamic parameters for the formation of single A–T and G–C pairs within a chain, the question can be asked: Does this permit calculation of the thermodynamic stability of a DNA of a given composition? Regrettably, this is not the case. The stability of a given DNA cannot be predicted simply by knowing the fraction of G–C pairs in the DNA.

Remarkably, if the effect of the nearest neighbors of each base pair is taken into account, a reasonable estimate of the thermodynamic stability of DNA can be obtained. This was discovered by studying many different complementary strands of short DNAs and looking for regularities (9,10). The rationale for this procedure lies in the importance of hydrogen bonding between each pair and the stacking interactions with its nearest neighbors. Ten nearest neighbor parameters suffice for determining the thermodynamic stability of a given DNA sequence, with one additional assumption, namely, the thermodynamic parameters for initiating DNA structure are different from those for adding hydrogen-bonded pairs to an existing chain. This takes into account that getting the two chains together and forming the first base pair is more difficult than adding base pairs to the double helix. Moreover,

TABLE 3-10. Thermodynamic Parameters for Determination of DNA Stability^a

DNA pair	ΔG° (kcal/mol)	ΔH° (kcal/mol)	ΔS° [cal/(mol K)]
5'-A-A	- 1.00	- 7.9	- 22.2
3'-T-T			
5'-A-T	- 0.88	- 7.2	- 20.4
3'-T-A			
5'-T-A	- 0.58	- 7.2	- 21.3
3'-A-T			
5'-A-C	- 1.44	- 8.4	- 22.4
3'-T-G			
5'-C-A	- 1.45	- 8.5	- 22.7
3'-G-T			
5'-A-G	- 1.28	- 7.8	- 21.0
3'-T-C			
5'-G-A	- 1.30	- 8.2	- 22.2
3'-C-T			
5'-C-G	- 2.17	- 10.6	- 27.2
3'-G-C			
5'-G-C	- 2.24	- 9.8	- 24.4
3'-C-G			
5'-C-C	- 1.84	- 8.0	- 19.9
3'-G-G			
Per GC at terminus	+ 0.98	+ 0.1	- 2.8
Per AT at terminus	+ 1.03	2.3	4.1

Source: Adapted from J. SantaLucia, Jr., SantaLucia, Jr. and D. Hicks, *Ann. Rev. Biophys. Biomolec. Structure* **33**, 415 (2004).

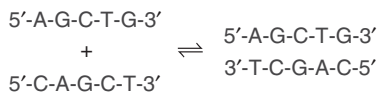
^apH 7.0, 1 M NaCl, 310 K.

the G-C hydrogen-bonded pair nucleates the double helix formation better than an A-T pair since it is more stable. (Only ten parameters are needed because some of the combinations of dinucleotides are redundant, for example, 5'GA3'/3'CT5' = 5'TC3'/3'AG5'.) In thermodynamic terms, this can be written as

$$\Delta G^\circ = \Delta G^\circ(\text{initiation}) + \sum \Delta G^\circ(\text{nearest neighbors}) \quad (3-20)$$

where the sum is over all pairs of nearest neighbor interactions. The thermodynamic parameters necessary to calculate the nearest neighbor interactions are given in Table 3-10, along with those characterizing the double helix initiation. The best fit of the data on model systems utilizes two additional parameters that are usually only small corrections; this refinement will not be considered here (cf. Ref. 9).

A simple example will illustrate how these data can be used. Consider the reaction



The sum of the nearest neighbor standard Gibbs energy changes associated with the formation of the five base pair DNA from Table 3-10 is

$$\begin{aligned} \sum \Delta G^\circ &= 5'-A-G + 5'-G-C + 5'-C-T + 5'-T-G \\ &\quad T-C-5' \quad C-G-5' \quad G-A-5' \quad A-C-5' \\ &= -1.28 - 2.24 - 1.28 - 1.45 = -6.25 \text{ kcal/mol} \end{aligned}$$

The standard Gibbs energy for the end effects is $0.98 + 1.03 = 2.01$ kcal/mol, so that the standard Gibbs energy change for formation of this DNA fragment is -4.24 kcal/mol. Similarly, $\Delta H^\circ = -31.5$ kcal/mol and $\Delta S^\circ = -87.8$ cal/(mol K). The temperature dependence of the standard Gibbs energy change can be calculated by assuming the standard enthalpy change is independent of temperature. This relatively simple procedure permits the thermal stability of any linear DNA to be calculated to a good approximation.

Knowledge of the thermal properties of DNA fragments is important both physiologically and practically. Knowing the stability of DNA obviously is of interest in understanding genetic replication. From a practical standpoint, knowing the stability of DNA fragments is important in planning cloning experiments. DNA probes must be used that will form stable duplexes with the target DNA. The temperature at which the duplex becomes stable can be estimated by calculating the melting temperature, T_m , that is, the temperature at which half of the strands are in the double helix conformation. The equilibrium constant for formation of the double helix duplex is

$$K = \frac{[D]}{[S]^2} \quad (3-21)$$

where D is the helical duplex and S is the single strand. If the total concentration of the oligonucleotide is C_0 , the concentration of single strands at T_m is $C_0/2$ and that of the duplex is $C_0/4$. If we insert these relationships into Eq. (3-21), we see that $K = 1/C_0$ at T_m . If we insert the relationship between the equilibrium constant and the standard Gibbs energy change, we obtain

$$\frac{\Delta G^\circ}{(RT_m)} = \ln C_0 \quad (3-22)$$

or

$$\frac{\Delta H^\circ}{(RT_m)} - \frac{\Delta S^\circ}{R} = \ln C_0$$

If Eq. (3-22) is solved for T_m , we obtain

$$T_m = \frac{\Delta H^\circ}{(\Delta S^\circ + R \ln C_0)} \quad (3-23)$$

For the case under discussion, if $C_0 = 0.1$ mM, $T_m = 297$ K.

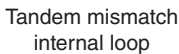
In this brief discussion, we have neglected the fact that DNA is a polyelectrolyte due to the negatively charged phosphodiester backbone. The polyelectrolyte nature of DNA means that the ionic environment, particularly positively charged ions, strongly influences the structure and behavior of DNA. Normally, a negatively charged polymer such as DNA would exist as a rod because of the charge repulsion. However, we know that DNA is packaged into a small volume in cells. This involves the twisting of the double helix, the formation of loops, etc. Interactions with metal ions and positively charged proteins are necessary for this packaging to occur. Interested readers should consult more complete descriptions of nucleic acids for information on this interesting subject (3,7).

3.9 RNA

In principle, the structure of RNA can be discussed exactly as the discussion of DNA. The principles are the same: hydrogen bonding between bases, stacking interactions, and hydrophobic interactions determine the structure. Of course, a ribose rather than a deoxyribose is present and uracil is substituted for thymine. In addition, several modified bases are commonly found in RNAs. Unfortunately, understanding and predicting RNA structures is more complex than for DNA. Ultimately, this is because there are several quite different biological functions for RNA and thus several quite different types of RNA. Generally, RNAs do not form intermolecular double helices. Instead, double helices are formed within an RNA molecule. These can be loops, hairpins, bulges, etc. Some idea of the diversity of structures that can be formed is shown in Figure 3-12, where an RNA molecule is shown, along with the different types of structures that can be formed.

As with DNA, the RNA structures can be predicted reasonably well by considering the nearest neighbor interactions. However, because of the diversity of the structures, the models are more complex. Nevertheless, quite reasonable RNA structures can be predicted using the thermodynamics derived from simple systems. We will not delve further into RNA structures here, except to say that the principles of model building have been developed sufficiently here so that the interested reader can proceed directly to current literature on this subject (11–14). As with DNA, metal ions play an important role in the biological packaging of RNA.

Finally, the interactions between RNA and DNA are of obvious physiological relevance. These have been studied quite extensively, but they are not nearly as well understood as the interactions between DNA fragments and within RNA molecules. Simple models are not yet available to calculate the properties of DNA–RNA structures. At this point, we will leave our discussion of thermodynamics of biological systems, although many more interesting examples could be discussed.



Copyright © 1997 Biopolymers. Reprinted by permission of John Wiley & Sons, Inc.

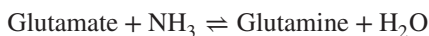
REFERENCES

1. P. Mitchell, *Nature* 191, 144 (1961).
2. T. G. Dewey and G. G. Hammes, *J. Biol. Chem.* 256, 8941 (1981).
3. C. R. Cantor and P. R. Schimmel, *Biophysical Chemistry*, Freeman, New York, 1980.
4. J. M. Berg, J. L. Tymoczko, and L. Stryer, *Biochemistry*, 7th edition, Freeman, New York, 2010.

5. C. Branden and J. Tooze, *Introduction to Protein Structure*, 2nd edition, Garland Publishing, New York, 1999.
6. W. Kauzmann, *Adv. Protein Chem.* 14, 1 (1959).
7. J. D. Watson, T. A. Baker, S. P. Bell, A. Gann, M. Levine, and R. Losick, *Molecular Biology of the Gene*, 7th edition, Benjamin/Cummings, Menlo Park, CA, 2013.
8. K. J. Breslau and J. M. Sturtevant, *Biophys. Chem.* 7, 205 (1977).
9. J. SantaLucia, Jr., H. T. Allawi, and P. A. Seneviratne, *Biochemistry* 35, 3555 (1996).
10. J. SantaLucia, Jr. and D. Hicks, *Ann. Rev. Biophys. Biomolec. Structure* 33, 415 (2004).
11. J. SantaLucia, Jr. and D. H. Turner, *Biopolymers* 44, 309 (1997).
12. T. Xia, J. SantaLucia, Jr., M. E. Burkard, R. Kierzek, S. J. Schroeder, X. Jiao, C. Cox, and D. H. Turner, *Biochemistry* 37, 14719 (1998).
13. J. F. Atkins, R. F. Gesteland and T. R. Cech *RNA Worlds: From Life's Origins to Diversity in Gene Regulation*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 2011
14. I. Yildirim, S. D. Kennedy, H. A. Stern, J. M. Hart, R. Kierzek, D. H. Turner, *J. Chem. Theory, Comput.* 8, 172 (2012).

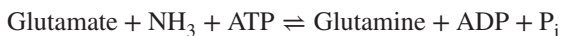
PROBLEMS

- 3-1. a.** Glutamine is an important biomolecule made from glutamate. Calculate the equilibrium constant for the reaction



Use the standard Gibbs energies of formation in Appendix 4 to obtain the standard Gibbs energy change for this reaction. Under physiological conditions, the concentration of ammonia is about 10 mM. Calculate the ratio $[\text{glutamine}]/[\text{glutamate}]$ at equilibrium. (By convention, the concentration of water is set equal to 1 since its concentration does not change significantly during the course of the reaction.)

- b.** Physiologically, glutamine is synthesized by coupling the hydrolysis of ATP to the aforementioned reaction:



Calculate ΔG° and the equilibrium constant for this reaction under standard conditions (see Appendix 4). Assume that NH_3 and P_i are maintained at about 10 mM and that $[\text{ATP}]/[\text{ADP}] = 1$. What is the ratio of glutamate to glutamine at equilibrium? What ratio is needed to convert glutamate to glutamine spontaneously, that is, to make $\Delta G < 0$?

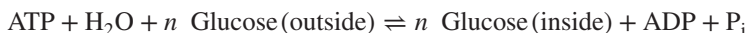
Do all calculations at 1 atm and 298 K. Although this is not the physiological temperature, the results are not significantly altered.

- 3-2.** Adipose tissues contain high levels of fructose. Fructose can enter the glycolytic pathway directly through the reaction



Assume that the standard Gibbs energy change for this reaction with the same standard state used in Table 3-1 is -17.0 kJ/mol . If fructose is substituted for glucose in “glycolysis,” what would the overall reaction be for the conversion of fructose to 2-glyceraldehyde, part one of glycolysis? What would the overall reaction be for the complete metabolic cycle? Calculate ΔG° and ΔG for these two reactions. Assume the concentration of fructose is 5.0 mM and the concentrations of the other metabolites are as in Table 3-2.

- 3-3.** Glucose is actively transported into red blood cells by coupling the transport with the hydrolysis of ATP. The overall reaction can be written as



If the ratio of $[\text{ATP}]/[\text{ADP}] = 1$ and $[\text{P}_i] = 10 \text{ mM}$, what is the concentration gradient, $[\text{glucose(inside)}]/[\text{glucose(outside)}]$, that is established at 298 K ? Calculate this ratio for $n = 1$, $n = 2$, and $n = n$. Does this suggest a method for determining the value of n ? Use the value of ΔG° for the hydrolysis of ATP in Appendix 5 for these calculations.

- 3-4.** Derive the equation for the temperature dependence of the standard Gibbs energy change for protein denaturation when ΔC_p is not equal to zero. As the starting temperature in the derivation, use the temperature at which $\Delta G^\circ = 0$.

Hint: The easiest way to proceed is to calculate the temperature dependence of ΔH° and ΔS° . These relationships can then be combined to give the temperature dependence of ΔG° . The parameters in the final equation should be ΔC_p , the temperature, T , the temperature at which $\Delta G^\circ = 0$, T_m , and the enthalpy change at the temperature where $\Delta G^\circ = 0$, $\Delta H_{T_m}^\circ$.

- 3-5.** Specific genes in DNA are often searched for by combining a radioactive oligonucleotide, O, with the DNA that is complementary to a sequence in the gene being sought. This reaction can be represented as



For such experiments, the concentration of the oligonucleotide is much greater than that of the specific DNA sequence. Assume that the probe is $5'\text{-GGGATCAG-3'}$.

- a.** Calculate the equilibrium constant at 310 K for the interaction of the probe with the complementary DNA sequence using the parameters in Table 3-9.

- b. Calculate the fraction of the DNA present as the double strand formed with the probe if the concentration of the probe is 1.0×10^{-4} M and the temperature is 310 K.
 - c. Find the melting temperature for this double strand if the concentration of the complementary DNA sequence is 1.0 nM. The melting temperature in this case is when $[\text{double strand}]/[\text{DNA}] = 1$.
- 3-6.** For an electrochemical cell (e.g., a battery), the reversible work is $-nF\varepsilon$ where n is the number of moles of electrons involved in the chemical reaction, F is the Faraday, and ε is the reversible voltage. (See Appendix 6). This relationship is useful for considering coupled oxidation–reduction reactions in biochemical systems.
- a. Use this relationship and Eq. (2-49) to derive an equation relating the Gibbs energy change for the reaction and the voltage of the cell. Your final equation should contain the standard Gibbs energy change for the reaction, the concentrations of the reactants, and the electrochemical voltage, as well as constants.
 - b. The voltage at equilibrium is usually designated as ε^0 . How is this related to the standard Gibbs energy change for the reaction? How might the equilibrium constant for a biochemical reaction be determined from voltage measurements? For the reaction



the voltage at equilibrium is -0.154 V at 298 K. Calculate the equilibrium constant for this reaction. ($F = 96,485$ coulomb/mol and $n = 2$ for this reaction.)

Thermodynamics Revisited

4.1 INTRODUCTION

Thus far thermodynamics has been developed using elementary calculus involving a single independent variable. This has been accomplished by specifying in the text that other variables are held constant when taking a differential. For example, $dE/dT = C_V$ at constant volume (Eq. (1–26)). A more general approach is possible by utilizing multivariable calculus. This provides not only a more exact mathematical formulation, but also relationships between thermodynamic variables that cannot be readily accessed within the restricted approach utilized thus far.

4.2 MATHEMATICAL TOOLS

If a function has multiple independent variables, that is, $y = f(x_1, x_2, x_3, \dots)$, then the function can be differentiated with respect to each variable, holding all other independent variables constant. This is called the partial derivative. For example, the partial derivative with respect to x_1 is written as $(\partial y / \partial x_1)_{x_2, x_3, \dots}$ where the subscripts outside the parentheses means these independent variables are held constant.

Consider the ideal gas law for one mole of gas, $P(V, T) = (RT)/V$. In this case,

$$\left(\frac{\partial P}{\partial T} \right)_V = \frac{R}{V} \quad (4-1)$$

$$\left(\frac{\partial P}{\partial V} \right)_T = -\frac{RT}{V^2} \quad (4-2)$$

A function can be differentiated sequentially with respect to each variable. For the case under discussion, the partial derivative of P with respect to T could be taken and then the partial derivative with respect to V could be taken, or vice versa. An important property of multiple partial differentiation is that the same result is obtained

regardless of the order of the differentiation. For the case being discussed, this means that

$$\left[\frac{\partial \left(\frac{\partial P}{\partial V} \right)_T}{\partial T} \right]_V = \left[\frac{\partial \left(\frac{\partial P}{\partial T} \right)_V}{\partial V} \right]_T \quad (4-3)$$

This property will prove useful in later discussions.

Finally, an exact differential can be written for a function in terms of the partial derivatives. For $y = f(x_1, x_2, x_3, \dots)$, the exact differential is

$$dy = \left(\frac{\partial y}{\partial x_1} \right)_{x_2, x_3, \dots} dx_1 + \left(\frac{\partial y}{\partial x_2} \right)_{x_1, x_3, \dots} dx_2 + \left(\frac{\partial y}{\partial x_3} \right)_{x_1, x_2, \dots} dx_3 + \dots \quad (4-4)$$

For $P = (RT)/V$,

$$dP = \left(\frac{\partial P}{\partial T} \right)_V dT + \left(\frac{\partial P}{\partial V} \right)_T dV \quad (4-5)$$

The exact differential can also be obtained by differentiating $P = (RT)/V$ with respect to each independent variable:

$$dP = \frac{R}{V} dT - \frac{RT}{V^2} dV \quad (4-6)$$

Examination of the aforementioned results shows that these two equations for the exact differential are equivalent. Most thermodynamic functions can be written as exact differentials in essence because their values are independent of the path. Quantities that are path dependent such as heat and work cannot be written as exact differentials.

Although we will not deal with integration of multiple variables, the principles are similar. For example,

$$\int \int f(x, y) dx dy$$

can be integrated first with respect to x with the appropriate integration limits and then with respect to y , again with defined integration limits. As with differentiation, the order of integration does not matter.

4.3 MAXWELL RELATIONS

We have previously shown the internal energy, E , can be expressed as a function of the entropy, S , and the volume, V :

$$dE = TdS - PdV \quad (4-7)$$

From the definition of the exact differential:

$$dE = \left(\frac{\partial E}{\partial S} \right)_V dS + \left(\frac{\partial E}{\partial V} \right)_S dV \quad (4-8)$$

Therefore, it follows that $(\partial E/\partial S)_V = T$ and $(\partial E/\partial V)_S = -P$. These equations can be differentiated further and since the order of differentiation does not matter:

$$\left[\frac{\partial \left(\frac{\partial E}{\partial S} \right)_V}{\partial V} \right]_S = \left[\frac{\partial \left(\frac{\partial E}{\partial V} \right)_S}{\partial S} \right]_V$$

Substitution of $(\partial E/\partial S)_V = T$ and $(\partial E/\partial V)_S = -P$ into the aforementioned equation gives the Maxwell relation:

$$\left(\frac{\partial T}{\partial V} \right)_S = - \left(\frac{\partial P}{\partial S} \right)_V \quad (4-9)$$

Three other Maxwell relations exist that can be obtained from the enthalpy, $H(S, P) = E + PV$, the Gibbs energy, $G(P, T) = H - TS$, and the Helmholtz energy, $A(V, T) = E - TS$. These are briefly derived as follows.

$$dH = dE + PdV + VdP = TdS + VdP = \left(\frac{\partial H}{\partial S} \right)_P dS + \left(\frac{\partial H}{\partial P} \right)_S dP$$

Therefore,

$$\left(\frac{\partial T}{\partial P} \right)_S = \left(\frac{\partial V}{\partial S} \right)_P \quad (4-10)$$

$$dG = dH - TdS - SdT = VdP - SdT = \left(\frac{\partial G}{\partial P} \right)_T dP + \left(\frac{\partial G}{\partial T} \right)_P dT$$

Therefore

$$\left(\frac{\partial V}{\partial T} \right)_P = - \left(\frac{\partial S}{\partial P} \right)_T \quad (4-11)$$

$$dA = dE - TdS - SdT = -PdV - SdT = \left(\frac{\partial A}{\partial V} \right)_T dV + \left(\frac{\partial A}{\partial T} \right)_V dT$$

Therefore

$$\left(\frac{\partial P}{\partial T} \right)_V = \left(\frac{\partial S}{\partial V} \right)_T \quad (4-12)$$

The Maxwell relations, Eqs. (4–9) through (4–12), permit the transformation between many different variables. For example, the energy can be expressed in terms of the variables P and T . Partial differentiation of Eq. (4-8) with respect to V at constant T gives:

$$\left(\frac{\partial E}{\partial V} \right)_T = \left(\frac{\partial E}{\partial S} \right)_V \left(\frac{\partial S}{\partial V} \right)_T + \left(\frac{\partial E}{\partial V} \right)_S = T \left(\frac{\partial S}{\partial V} \right)_T - P$$

Insertion of Eq. (4-12) results in

$$\left(\frac{\partial E}{\partial V}\right)_T = T\left(\frac{\partial P}{\partial T}\right)_V - P \quad (4-13)$$

Eq. (4-13) is called a thermodynamic equation of state as it expresses the energy derivative with respect to V in terms of the variables P and T .

As an example of the use of multivariable calculus, the temperature dependence of the Gibbs energy is derived (analogous to Eq. (2-43) for ΔG). From the total differential for G given earlier,

$$\left(\frac{\partial G}{\partial T}\right)_P = -S \quad (4-14)$$

Since S is positive, G decreases when the temperature is raised at constant pressure. The temperature dependence of G can also be expressed in terms of the enthalpy by substituting for S using the definition of G , namely $S = (H - G)/T$:

$$\left(\frac{\partial G}{\partial T}\right)_P - \frac{G}{T} = -\frac{H}{T^2}$$

Also note by the rules of calculus,

$$\left[\frac{\partial\left(\frac{G}{T}\right)}{\partial T}\right]_P = \left(\frac{1}{T}\right)\left(\frac{\partial G}{\partial T}\right)_P - \frac{G}{T^2}$$

Substitution into the preceding equation gives the Gibbs-Helmholtz equation, the equivalent of Eq. (2-43):

$$\left[\frac{\partial\left(\frac{G}{T}\right)}{\partial T}\right]_P = -\frac{H}{T^2} \quad (4-15)$$

4.4 CHEMICAL POTENTIAL

Thus far, the implicit assumption has been made that the thermodynamic system contains only a single component. However, systems of interest to chemists and biologists typically contain many components. In an open system, that is, one in which the composition may vary, the use of multivariable calculus allows for an easy extension of our previous formulations. For example, the exact differential for the Gibbs energy as a function of T , P , and the number of moles of each substance, n_i , $G(T, P, n_1, n_2, n_3, \dots)$, can be written as

$$dG = -SdT + VdP + \left(\frac{\partial G}{\partial n_1}\right)_{T,P,n_2,n_3,\dots} dn_1 + \left(\frac{\partial G}{\partial n_2}\right)_{T,P,n_1,n_3,\dots} dn_2 + \dots \quad (4-16)$$

The chemical potential of each component is defined as the partial derivative of the Gibbs energy of each component with respect to the number of moles of that component, with all other independent variables being held constant. Thus, the chemical potential for n_1 , μ_1 , is

$$\mu_1 = \left(\frac{\partial G}{\partial n_1} \right)_{T,P,n_2,n_3,\dots} \quad (4-17)$$

The chemical potential, μ , has been briefly introduced in the discussion of the relationship between the molar Gibbs energy and the activity (Eq. (2-33)). For rigorous thermodynamic treatments of solutions (and mixtures in general), it is the function of choice. For a single pure component, the chemical potential is the Gibbs energy per mole. For a mixture, the Gibbs energy change will depend on the other components present so that the other components and their concentrations must be specified.

Eq. (4-16) can then be conveniently expressed as:

$$\begin{aligned} dG &= -SdT + VdP + \mu_1 dn_1 + \mu_2 dn_2 + \mu_3 dn_3 + \dots \\ &= -SdT + VdP + \sum \mu_i dn_i \end{aligned} \quad (4-18)$$

where the sum is over all of the components present.

In order to apply the chemical potential to real systems, the relationship between the chemical potential and the concentration of each component must be established. The starting point for the development of this relationship is to consider the chemical potential for mixtures of liquids. The chemical potential of an ideal gas (Gibbs energy/mole) is (Eq. (2-28)) $\mu_i = \mu_i^\circ + RT \ln (P_i/P_0)$. Since the chemical potential of the liquid and vapor must be equal at equilibrium, the chemical potential of the pure liquid A is

$$\mu_A^\circ = \mu_{A\text{gas}}^\circ + RT \ln \left(\frac{P_{A0}}{P_0} \right)$$

where P_{A0} is the vapor pressure of the pure liquid. For a mixture of liquids, the vapor pressure is P_A and the chemical potential of the liquid is

$$\mu_A = \mu_{A\text{gas}}^\circ + RT \ln \left(\frac{P_A}{P_0} \right)$$

Combining these two equations gives the chemical potential for A in a mixture:

$$\mu_A = \mu_A^\circ + RT \ln \left(\frac{P_A}{P_{A0}} \right) \quad (4-19)$$

The relationship between the ratio of pressures and the mole fraction, X_A , was established empirically by Francois Raoult, who found that for many binary mixtures of liquids

$$P_A = X_A P_{A0} \quad (4-20)$$

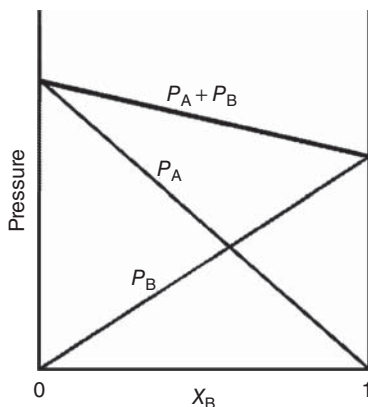


FIGURE 4-1. Plot of the vapor pressure of liquids A and B and the total vapor pressure versus the mole fraction of liquid B for a binary mixture of liquids that obey Raoult's law.

The chemical potential can then be written as

$$\mu_A = \mu_A^0 + RT \ln X_A \quad (4-21)$$

Mixtures of liquids that obey Raoult's law are called "ideal" solutions. This law is obeyed quite well for very similar mixtures such as various benzene derivatives, but large deviations are found for dissimilar liquids. An illustrative plot of the vapor pressure of each liquid and the total vapor pressure versus the mole fraction for a binary mixture of two liquids obeying Raoult's law is shown in Figure 4-1.

Deviations from ideal solution behavior are taken into account by defining the chemical potential in terms of the thermodynamic activity, a ,

$$\mu_i = \mu_i^0 + RT \ln a_i \quad (4-22)$$

where μ_i^0 is the standard chemical potential, that is, the chemical potential when the activity, a_i , is equal to 1. The activity can be written as $\gamma_i X_i$, where γ_i is the activity coefficient, an empirical parameter that takes into account deviations from ideality.

Two limiting cases merit special consideration. If one component of the mixture is present in a much larger concentration than all other components so that it becomes the solvent, then Raoult's law is obeyed quite well. At the other extreme, in a binary mixture where one liquid is the solvent, the other liquid is present in very low concentrations and obeys Henry's law, $P_B = K_B X_B$ where K_B is an empirical constant that is not the equilibrium vapor pressure of pure B. An ideal solution for the solute is a dilute solution that obeys Henry's law and a chemical potential can be written utilizing Henry's law.

In order for the chemical potential to be useful in biological systems, solute chemical potentials must be related to concentration scales normally used such as molality and molarity. To make this connection, we note that for dilute solutions, the mole

fraction of the solute, X_B , can be approximated by n_B/n_A where n_B and n_A are the number of moles of solute and solvent, respectively. If the molality concentration scale is used, the number of moles of solute in a kg of solvent is m_B and the number of moles of the solvent is $1\text{kg}/\text{MW}_A$ where MW_A is the molecular weight of the solvent A. Therefore, $X_B = m_B \text{MW}_A$. Since the molality is related to the mole fraction by a constant, the constant can be lumped into the standard chemical potential to give $\mu_i = \mu_i^\circ + RT \ln m_i$ for an ideal solution. This equation can be generalized to nonideal solutions by use of the concepts of thermodynamic activity and activity coefficients to give

$$\mu_i = \mu_i^\circ + RT \ln a_i = \mu_i^\circ + RT \ln \gamma_i m_i \quad (4-23)$$

Since molality has dimensions of moles per kg of solvent, Eq. (4-23) would seem to imply that the natural logarithm of a number with dimensions is being taken. This, however, is not the case as the molality in this equation is implicitly divided by the standard state molality usually taken as 1 m. Despite this mathematical ambiguity, the units of the concentration should always be specified when using Eq. (4-23). Stipulation of the concentration units automatically specifies the concentration units of the standard state. However, bear in mind that you are really dealing with a dimensionless number.

The aforementioned discussion has utilized molality as the unit for concentration. A similar argument could be advanced for utilization of molarity, which is more universally used by biologists. From a practical standpoint, the use of molarity has a significant drawback because it is dependent on temperature and pressure as the volume of solvent varies with temperature and pressure. For this reason, the molality is a better measure of concentration to use in the aforementioned equations. Fortunately, for dilute aqueous solutions at room temperature, molarity and molality are numerically approximately equal.

The definition of standard state becomes somewhat confusing with the aforementioned formulation. For Eqs. (4-22) and (4-23), the standard state is the *hypothetical* state of unit activity. For Eq. (4-23), a standard state must be chosen where γ approaches 1. This occurs when no interactions exist between solutes, that is, at infinite dilution. In this case, the standard state is a *hypothetical* molality (or molarity) of unity where $\gamma = 1$. From a practical standpoint, the designation of standard state does not have any effect on the use of the chemical potential as long as it is clearly defined.

4.5 PARTIAL MOLAR QUANTITIES

We now digress to discuss the concept of partial molar quantities, which are of great utility when considering mixtures. As an example, consider the volume. At constant temperature and pressure, changes in volume, dV , for a mixture can be written as

$$dV = \left(\frac{\partial V}{\partial n_1} \right)_{T,P,n_2,n_3,\dots} dn_1 + \left(\frac{\partial V}{\partial n_2} \right)_{T,P,n_1,n_3,\dots} dn_2 + \dots \quad (4-24)$$

The partial molar volume for component n_1 , $\bar{V}_1$, is defined as

$$\bar{V}_1 = \left(\frac{\partial V}{\partial n_1} \right)_{T,P,n_2,n_3,\dots} \quad (4-25)$$

Note that the partial molar volume is a function of the temperature, the pressure, and the composition of the mixture. For a single component, the meaning of the partial molar volume is clear: the partial molar volume is a constant at constant temperature and pressure. For example, for water it is 18 cc at 298 K and 1 atm. However, if water is mixed with another solvent, the partial molar volume of water will depend on the amount of the second component. For example, if a small amount of water is added to a large amount of pure ethanol, the partial molar volume of water is 14 cc. This is because the water molecules tend to pack closer together to avoid contact with the hydrocarbon portion of ethanol.

For a two component system, the total volume can be written as

$$V = \bar{V}_1 n_1 + \bar{V}_2 n_2 \quad (4-26)$$

where the partial molar volumes are those at the specific composition n_1 and n_2 . Partial molar volume can be measured experimentally. If the total volume is measured as a function of the composition, the slope dV/dn at any given composition is the partial molar volume. This simple experiment works, but more accurate methods of data analysis are required for determining the partial molar volume.

From these considerations, the chemical potential can be seen to be the partial molar Gibbs energy (Eq. (4-17)). By analogy from the results obtained when considering the volume, for a two component system at constant temperature and pressure:

$$G = n_1 \mu_1 + n_2 \mu_2$$

Differentiating the above equation results in

$$dG = n_1 d\mu_1 + n_2 d\mu_2 + \mu_1 dn_1 + \mu_2 dn_2$$

However, from Eq. (4-18), $dG = \mu_1 dn_1 + \mu_2 dn_2$. In order for both of these results to be correct, $n_1 d\mu_1 + n_2 d\mu_2 = 0$. This can be generalized to give the Gibbs-Duhem equation:

$$\sum n_i d\mu_i = 0 \quad (4-27)$$

The sum extends over all of the components in the system.

The significance of this result is that the chemical potentials in a system are not independent. A change in one chemical potential must be balanced by a change or changes of the other chemical potentials. For example, in the two component system if one of the chemical potentials increases, the other must decrease. The Gibbs-Duhem equation is of great utility in dealing with mixtures.

4.6 OSMOTIC PRESSURE

Osmotic pressure provides an example of the utility of the chemical potential. Cell membranes are semipermeable; that is, they allow some molecules and ions to pass freely through them, but do not permit others to pass. Typically large molecules such as proteins and nucleic acids are confined inside the cell, whereas small ions and molecules pass through the membrane. This is a gross simplification as biological membranes have specialized transport and signaling mechanisms that make transport in and out of the cell very complex. At equilibrium, the chemical potentials of the various components should be equal on both sides of the membrane. For a completely permeable membrane, this is achieved by transport through the membrane until the concentrations (or more precisely activities) of all components are equal on both sides of the membrane. However, for a semipermeable membrane, this cannot happen so equilibrium must be achieved in a different way. Since the temperature is constant, equilibrium is achieved by establishing a pressure difference across the membrane. This pressure difference is called the *osmotic pressure*. This pressure difference is created by solvent flowing into the solution containing the substance that cannot pass through the membrane. This process is called *osmosis*.

The osmotic pressure can be calculated using the thermodynamic principles that have been discussed. A simplified experimental setup demonstrating the phenomenon of osmotic pressure is shown in Figure 4-2. In this experiment, a constant pressure, P , is applied to the pure solvent. In order to keep the solutions on each side of the semipermeable membrane at the same level, an additional pressure, the osmotic pressure, Π , must be applied to the solution. The actual experimental equipment for measuring osmotic pressure, an osmometer, is more complex and smaller and measures differences in solution level on the two sides of the semipermeable membrane.

The value of Π for a dilute solution of a macromolecule will now be derived. At equilibrium the chemical potential of the solvent on each side of the membrane must

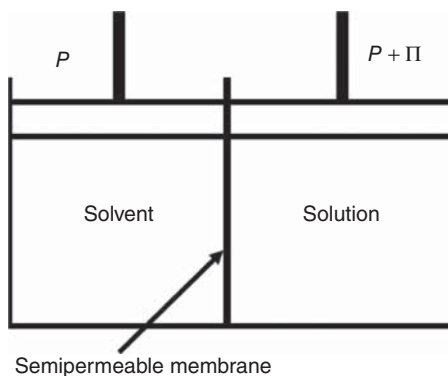


FIGURE 4-2. Schematic representation of the concept of osmotic pressure. The solution contains a substance that cannot pass through the membrane. In order for the chemical potential to be equal for the solvent and solution, an additional pressure, Π , must be applied to the solution.

be the same. If we consider a slight deviation from equilibrium, $d\mu$, then

$$d\mu = \left(\frac{\partial \mu}{\partial P} \right)_{X_S} dP + \left(\frac{\partial \mu}{\partial X_S} \right)_P dX_S \quad (4-28)$$

In this equation, P is the pressure and X_S is the mole fraction of solvent.

At equilibrium, $d\mu = 0$. Furthermore, since μ is the partial molar Gibbs energy, $(\partial \mu / \partial P)_{X_S}$ is equal to the partial molar volume of the solvent, $\bar{V}$, that is, the volume per mole of solvent. (This follows from the equation preceding Eq. (4-11).) In addition, $(\partial \mu / \partial X_S)_P = RT/X_S$ (Eq. (4-21)). Rearranging Eq. (4-28) for a system at equilibrium gives

$$\bar{V} dP = - \left(\frac{RT}{X_S} \right) dX_S \quad (4-29)$$

This equation can be integrated as follows:

$$\int_P^{P+\Pi} \bar{V} dP = -RT \int_1^{X_S} \frac{dX_S}{X_S} \quad (4-30)$$

$$\Pi \bar{V} = -RT \ln X_S = -RT \ln(1 - X_M)$$

where X_M is the mole fraction of the macromolecule. This derivation assumes a very dilute solution so that X_M is much less than 1. In this case, $\ln(1 - X_M) \approx -X_M$. Furthermore, $X_M = n_M/(n_M + n_S) \approx n_M/n_S$ where the n_i represent the moles of each component. Putting this all together gives

$$\Pi \bar{V} n_S = RT n_M \quad \text{or} \quad \Pi = RTc \quad (4-31)$$

The final rearrangement arises because $\bar{V} n_S$ is the total volume of solvent and the concentration of the macromolecule, c , is equal to n_M/V . This simple relationship is only valid for dilute solutions.

Eq. (4-31) permits the osmotic pressure to be calculated. Alternatively, if the osmotic pressure is measured, the concentration of the macromolecule can be determined. In fact, osmotic pressure measurements can be used to measure the unknown molecular weight of a macromolecule since $c = [g/(MW)]/V_{\text{solvent}}$ where g is the grams of solute, MW is the molecular weight, and V_{solvent} is the volume of solvent in liters. This is, in fact, the primary use of osmotic pressure measurements in the laboratory. However, Eq. (4-31) is only valid at very low concentrations so that the data analysis utilizes a more complex equation, which reduces to Eq. (4-31) as the macromolecule concentration approaches zero. Osmotic pressures are not large relative to atmospheric pressure: for a 10^{-4} M solution at room temperature Π is about 2.5×10^{-3} atm. This is a significant pressure for a cell wall, however, and can be measured quite accurately in laboratory experiments.

As an example of the role of osmotic pressure in biology, consider the immersion of red blood cells in hypertonic, isotonic, and hypotonic solutions. Hypertonic means

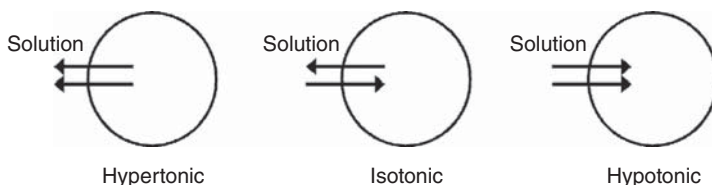


FIGURE 4-3. Schematic representation of red blood cells in hypertonic, isotonic, and hypotonic solutions. In hypertonic solutions, some of the solution inside the cell passes out of the cell and the cell shrinks. In hypotonic solutions, solution from the exterior of the cell passes into the cell, causing the cell to expand. In isotonic solutions, the flow of material between the two sides of the membrane is exactly balanced.

the concentration of solutes that cannot pass through the red cell membrane are higher on the outside than on the inside. Hypotonic is the converse, and isotonic means the concentrations are the same outside and inside the cell. For hypertonic and hypotonic solutions, an osmotic pressure exists across the membrane. In the case of hypertonic solutions, water flows out from the cell to equalize the pressure and the cell shrinks, whereas for hypotonic solutions, water flows into the cell, which expands and could even burst. In fact, dilution of cells into solvent is a method for lysing cells. For isotonic solutions, there is no osmotic pressure so that the flow of water into the cell balances the flow out of the cell. This situation is shown schematically in Figure 4-3.

4.7 CHEMICAL EQUILIBRIA

The subject of chemical equilibrium has already been discussed in Chapter 2. However, we now examine chemical equilibria using the chemical potential as a tool. In addition, the concepts of thermodynamic activity, activity coefficients, and standard states are considered in more detail.

As a starting point, consider a chemical reaction such as



At constant temperature and pressure:

$$dG = \mu_A dn_A + \mu_B dn_B + \mu_C dn_C + \mu_D dn_D \quad (4-33)$$

Furthermore, the stoichiometry requires that

$$\frac{dn_A}{a} = \frac{dn_B}{b} = -\frac{dn_C}{c} = -\frac{dn_D}{d} = -dx$$

Therefore:

$$\begin{aligned}
 dG &= (-a\mu_A - b\mu_B + c\mu_C + d\mu_D)dx \\
 dG &= -a(\mu_A^\circ + RT \ln a_A) - b(\mu_B^\circ + RT \ln a_B) + c(\mu_C^\circ + RT \ln a_C) \\
 &\quad + d(\mu_D^\circ + RT \ln a_D) \\
 dG &= \Delta\mu^\circ + RT \ln \left(\frac{a_D^d a_C^c}{a_B^b a_A^a} \right) \quad \text{with } \Delta\mu^\circ = c\mu_C^\circ + d\mu_D^\circ - a\mu_A^\circ - b\mu_B^\circ
 \end{aligned}$$

At equilibrium at constant T and P , $dG = 0$ and

$$\Delta\mu^\circ = -RT \ln \left(\frac{a_D^d a_C^c}{a_B^b a_A^a} \right) = -RT \ln K \quad (4-34)$$

Equation (4-34) is equivalent to the more commonly cited equation,

$$\Delta G^\circ = -RT \ln \left(\frac{a_D^d a_C^c}{a_B^b a_A^a} \right) = -RT \ln K \quad (4-35)$$

These equations are exact in that the standard state is precisely defined as the state where all activities are equal to unity, and the standard Gibbs energy change, ΔG° , is precisely the change in Gibbs energy when all reactants have an activity of unity. In these equations, the activities are the activities at equilibrium. However, from a practical standpoint, these equations are not very useful as concentrations, not activities, of reactants are measured. This problem is addressed by equating the activity to the product of an activity coefficient multiplied by the concentration, γ_c :

$$\Delta G^\circ = -RT \ln \left(\frac{\gamma_D^d \gamma_C^c}{\gamma_B^b \gamma_A^a} \right) - RT \ln \left(\frac{c_D^d c_C^c}{c_B^b c_A^a} \right) = -RT \ln K \quad (4-36)$$

Any convenient measure of concentration can be used, as long as the system used is self-consistent. However, the numerical values of K and ΔG° depend on the concentration scale chosen. As discussed earlier, molality is preferable to molarity. The dimensionless activity coefficients are dependent on the specific composition of the solution. If the composition is held constant, e.g., constant buffer concentration and pH in biological systems, then the usual assumption is that the activity coefficient ratio is constant and Eq. (4-36) becomes:

$$\Delta G^\circ + RT \ln \left(\frac{\gamma_D^d \gamma_C^c}{\gamma_B^b \gamma_A^a} \right) = \Delta G^{\circ'} = -RT \ln \left(\frac{c_D^d c_C^c}{c_B^b c_A^a} \right) \quad (4-37)$$

where the effective change in the standard Gibbs energy, $\Delta G^{\circ'}$, is defined by Eq. (4-37). However, some ambiguity exists with regard to the standard state. The

standard state can be considered to be the *hypothetical* state where the concentrations are unity and the activity coefficients approach unity (infinite dilution). As previously discussed, bear in mind that the concentrations in these equations are divided by the standard state concentration of unity so that the ratio of concentrations is dimensionless. In some case, activity coefficients can be measured and/or calculated. However, most experimentalists choose to ignore the activity coefficients and keep the solution environment the same during the course of their experiments. In the next section, ionic solutions are discussed. This is one of the few cases where a relatively simply calculation of the activity coefficient is possible.

4.8 IONIC SOLUTIONS

Ionic solutions require special consideration in the application of thermodynamics. If a strong electrolyte is dissolved in water, it forms positive and negative ions. For example, you might simply write down the chemical potential for a positive ion in terms of its concentration. However, this makes no sense because positive ions cannot exist in isolation. Their charge must be exactly balanced by those of negatively charged ions. A reasonable way to proceed is to consider the sum of the chemical potentials for the positive and negative ions:

$$\begin{aligned}\mu_+ + \mu_- &= \mu^{o+} + \mu^{o-} + RT \ln a_+ + RT \ln a_- \\ &= \mu^{o+} + \mu^{o-} + RT \ln \gamma_+ + RT \ln \gamma_- + RT \ln m_+ + RT \ln m_- \quad (4-38)\end{aligned}$$

The molality, m , has been selected as the concentration scale. The individual molalities can be readily calculated, but the two activity coefficients cannot be separated, as they always appear as a product, i.e., $\ln [\gamma_+ \gamma_-]$. This problem can be addressed by defining a mean activity coefficient, $\gamma_{\pm}$,

$$\gamma_{\pm}^s = \gamma_+^c \gamma_-^d \quad (4-39)$$

In this equation, $s = c + d$, where c and d are the stoichiometric coefficients for the cations and anions, respectively, of the strong electrolyte. For example, they both are 1 for NaCl whereas for CaCl_2 , $c = 1$ and $d = 2$. With this definition of the mean activity coefficient, the activity coefficient term in Eq. (4-34) becomes $sRT \ln \gamma_{\pm}$. Thus the determination of the mean activity coefficient is seen to require calculation of the energy term in the chemical potential that determines $sRT \ln \gamma_{\pm}$. Mean molalities and activities can be calculated analogous to the mean activity coefficient:

$$\begin{aligned}m_{\pm}^s &= m_+^c m_-^d \\ a_{\pm}^s &= a_+^c a_-^d\end{aligned}$$

The theory for calculation of the mean activity coefficient was developed in the late 1920s by Peter Debye and Erich Hückel, a graduate student at the time. Amazingly,

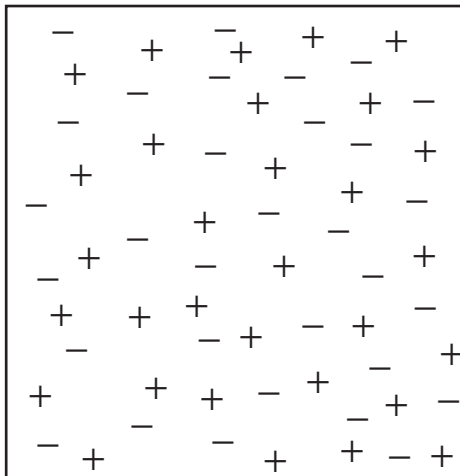


FIGURE 4-4. Pictorial representation of the distribution of positive and negative ions in one-one strong electrolyte solutions. Positive ions tend to be surrounded by negative ions and vice versa. In the Debye–Hückel theory, the ions are assumed to be point charges.

the Debye–Hückel theory is still used today, with a few embellishments. The premise is that interactions between charged ions are the only source of nonideality that need to be considered. This assumption is valid because the energies associated with these interactions are much greater than with any other type of interaction between ions in solution such as short range van der Waals interactions. From a thermodynamic standpoint, the addition to the chemical potential is the work required to create the ionic state from an uncharged state.

The Debye–Hückel model treats an electrolyte solution as a collection of point charges, as depicted schematically in Figure 4-4. Note that plus charges tend to be surrounded by minus charges and vice versa. Thus each ion has an “ionic atmosphere” surrounding it. The Coulomb potential, Ψ_i at a distant r from a charge $z_i e$ is

$$\Psi_i = \left(\frac{z_i e}{4\pi\epsilon} \right) \left(\frac{1}{r} \right) = \frac{Z_i}{r} \quad (4-40)$$

Here ϵ is the permittivity of the medium and Z_i is defined by Eq. (4-40). For a pair of charges, the potential energy of the system is Ψ_i multiplied by the second charge. When multiple charged ions are present in solution, the potential of a single ion is

$$\Psi_i = \left(\frac{Z_i}{r} \right) e^{-\kappa/r} \quad (4-41)$$

The exponential appearing in this equation is due to the “screening” of the single ion by the other ions in the system. The quantity κ is a length parameter that was derived by Debye and Hückel and can be thought of as the effective radius of the ionic atmosphere.

For one-one electrolytes at room temperature in water,

$$\kappa(\text{nm}) = \frac{0.305}{\sqrt{I}} \quad (4-42)$$

In this equation, I is the ionic strength calculated using the molal concentration scale. The ionic strength is defined by

$$I = \left(\frac{1}{2}\right) \sum z_i^2 m_i \quad (4-43)$$

The sum is over all ions present in solution, not only those involved in the mean activity coefficient under consideration. The ionic strength is actually dimensionless as each concentration should be divided by its standard concentration. However, as noted previously, assigning units to the ionic strength provides an unambiguous designation of the concentration scale being used. For an ionic strength of $0.01(m)$, $\kappa = 3.0 \text{ nm}$. As would be expected, κ decreases as the concentration of ions increases, that is, more “screening” takes place.

Debye and Hückel were able to obtain an explicit expression for the addition to the chemical potential needed to determine the mean activity coefficient. For water at 298 K, 1 atm, the final result is

$$\log \gamma_{\pm} = -0.509 |z_+ z_-| \sqrt{I} \quad (4-44)$$

Because of assumptions made in its derivation, Eq. (4-43) and Eq. (4-44) are valid only at very low concentrations of salts, typically less than 10^{-2} M . The vertical lines on the sides of the product of the two charges indicate the absolute value of the product should be used. This equation is the limiting law, that is, the law as the salt concentrations approach zero. One of the assumptions in the derivation is that the ions are point charges, whereas real ions have a finite size. If this approximation is removed, Eq. (4-44) becomes

$$\log \gamma_{\pm} = \frac{[-0.509 |z_+ z_-| \sqrt{I}]}{[1 + B \sqrt{I}]} \quad (4-45)$$

Here B is a constant related to the effective radius of an ion: in practice it is an adjustable parameter that can be used to fit experimental data. At low ionic strengths, the second term in the denominator becomes much less than one, and Eq (4-45) reduces to Eq. (4-44). Attempts to extend Eq. (4-45) to higher concentration have been made. For example, a linear term, CI , has been added, with C being an adjustable parameter. However, the beauty of the Debye–Hückel theory is its relative simplicity, and the model itself is an inadequate representation of the real system for concentrated solutions. A more complex model is required for real solutions, but thus far none of the models developed have provided a widely accepted replacement for the Debye–Hückel theory.

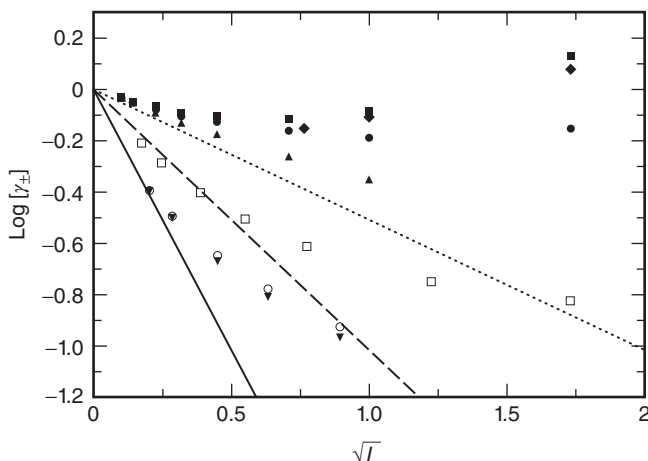


FIGURE 4-5. Plot of the logarithm of the mean ionic activity coefficient for various electrolytes versus the square root of the ionic strength (standard state is 1 m). The top four symbols are for NaCl, HCl, LiCl, and KNO₃ from top to bottom. The open squares are for H₂SO₄, and the bottom two symbols are for MgSO₄ and CuSO₄. The lines are those predicted by the Debye–Hückel theory for a +1,–1 electrolyte, a +1,–2 electrolyte, and a +2,–2 electrolyte, again from top to bottom (Eq. (4-44)). The data are from N. G. Lewis and M. Randall M. J. *Am. Chem. Soc.* **43**, 1112 (1921). Figure courtesy of Professor T. G. Oas, Duke University. Reproduced with permission.

Many tests of the Debye–Hückel theory have been carried out. For example, mean ionic activity coefficients can be measured experimentally, primarily through electrochemical experiments. An introduction to electrochemistry is given in Appendix 6. Although these measurements are not discussed in this treatise, some of the results obtained are shown in Figure 4-5, where the logarithm of the mean ionic activity coefficient is plotted versus the ionic strength for some selected salt solutions, including +1,–1 electrolytes, a +1,–2 electrolyte, and +2,–2 electrolytes. The limiting slopes predicted by the Debye–Hückel theory are also shown, and the data approach these slopes at low salt concentrations. However, significant deviations are seen at concentrations above about 10^{-2} (molal) ionic strengths. The deviations occur at lower salt concentrations for higher charged ions. This is predicted by the Debye–Hückel theory and is due to the approximations used to obtain Eq. (4-44).

Another test of the Debye–Hückel theory can be made by measuring the salt dependence of acid ionization constants. The dissociation of a weak acid can be written as



The equilibrium constant for this reaction can be obtained by a variety of methods, for example a pH titration. The observed equilibrium constant is

$$K_{\text{obs}} = \frac{[\text{H}^+][\text{A}^-]}{[\text{HA}]} \quad (4-47)$$

The thermodynamic equilibrium constant is

$$K_0 = \frac{\gamma_{\pm}^2 [\text{H}^+][\text{A}^-]}{[\text{HA}]} \quad (4-48)$$

The activity coefficient of the uncharged species, HA, has been set equal to 1 in Eq. (4-48). If the logarithm of Eq. (4-48) is taken and Eq. (4-44) is substituted for the mean ionic activity coefficients, we obtain

$$\log K_0 = 2 \log \gamma_{\pm} + \log K_{\text{obs}} = -1.018 \sqrt{I} + \log K_{\text{obs}} \quad (4-49)$$

Thus, a plot of $\log K_{\text{obs}}$ versus the square root of the ionic strength should be a straight line with a slope of 1.018 in the limit of low salt concentrations and with an intercept of $\log K_0$. As a practical matter, since the ionization constant varies with salt concentration, so will the pH of solutions of weak acids and bases.

This chapter has dealt with some of the more important aspects of advanced thermodynamics but is far from complete. Interested readers should consult the references at the end of the chapter for more complete treatments (cf. Refs. 1–4).

REFERENCES

1. R. A. Alberty, *Thermodynamics of Biochemical Reactions*, Wiley and Sons, Hoboken, NJ, 2003.
2. P. Atkins and J. de Paul, *Physical Chemistry*, 9th edition, W. H. Freeman and Co., New York, NY, 2010.
3. I. M. Klotz and R. M. Rosenberg, *Chemical thermodynamics*, 2nd edition, Benjamin, New York, NY, 1972.
4. G.N. Lewis and M. Randall, revised by K. S. Pitzer L. Brewer, *Thermodynamics*, McGraw Hill, New York, NY, 1961.

PROBLEMS

- 4-1. Derive an expression for the temperature dependence of the Helmholtz energy ($A = E - ST$), $[\partial(A/T)/\partial T]_V$, in terms of the temperature and internal energy, E , and the temperature.
- 4-2. Another thermodynamic equation of state is

$$\left(\frac{\partial H}{\partial P}\right)_T = V - T\left(\frac{\partial V}{\partial T}\right)_P$$

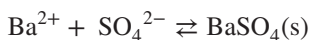
Use the Maxwell relations to derive this equation.

- 4-3. The partial molar volume of pure water and ethanol at 298 K and 1 atm are 18.0 cc and 58.0 cc, respectively. At a mole fraction of 0.500, the partial molar volumes are 16.9 cc and 57.4 cc, respectively.

- a. Calculate the volume change if 3.00 moles of water are mixed with 3.00 moles of ethanol.
- b. Derive an expression for the volume change when n_1 moles of ethanol are mixed with n_2 moles of water with a final mole fraction of ethanol of X_1 .
- 4-4. Some osmotic pressure data are given in the following table. The solvent is water and the temperature is 25.0° C. Obtain a molecular weight for the unknown substance from these data. You will note that a different molecular weight is obtained for each concentration of the unknown. How do you explain this? Can you suggest a method for obtaining the true molecular weight. Estimate the true molecular weight from these data.

Concentration (g/L)	Osmotic Pressure (atm)
0.500	0.000341
1.00	0.000678
1.50	0.00101
2.00	0.00134

- 4-5. CaCl_2 is a strong electrolyte that dissociates into calcium and chloride ions in water.
- a. Define the mean ionic activity coefficient for this salt.
- b. Write down an expression for the chemical potential of CaCl_2 in water in terms of the standard chemical potentials, the mean ionic activity coefficient, and the molality of CaCl_2 .
- c. Calculate the ionic strength of a 0.01 m CaCl_2 solution in water.
- d. Calculate the mean ionic activity coefficient for a 0.01 m aqueous solution of CaCl_2 at 25° C.
- e. NaCl is added to the CaCl_2 solution at a concentration of 0.01 m. Calculate the ionic strength of the resulting solution.
- f. Calculate the mean ionic activity coefficient for CaCl_2 at 25° C for the solution in e.
- 4-6. The solubility of salts in water is dependent on the ionic strength. For example, BaSO_4 is not very soluble in water. The equilibrium between the ions and the solid can be written as:



Derive an expression for the ionic strength dependence of the solubility product $K_{\text{sol}} = (\text{Ba}^{2+})(\text{SO}_4^{2-})$. The activity of the precipitated solid can be assumed to be 1. Based on this analysis, would you expect the solubility to increase or decrease as the ionic strength is increased.

CHEMICAL KINETICS

Principles of Chemical Kinetics

5.1 INTRODUCTION

Thermodynamics tells us what changes in state can occur, that is, the relative stability of states. For chemical reactions, it tells us what reactions can occur spontaneously. However, thermodynamics does not tell us the time scale for changes in state or how the changes in state occur; it is concerned only with the differences in the initial and final states. In terms of chemical reactions, it does not tell us *how* the reaction occurs, in other words, the molecular interactions that take place as a reaction occurs. For biological reactions, the rates are critical for the survival of the organism, and a primary interest of modern biology is the molecular events that lead to reaction. The study of the rates and mechanisms of chemical reactions is the domain of *chemical kinetics*. Thermodynamics provides no intrinsic information about mechanisms.

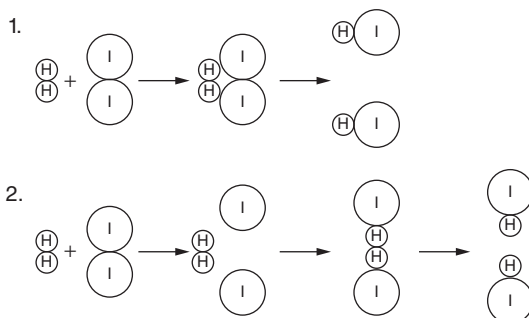
Many examples exist with regard to the importance of the time scale for chemical reactions. For those of you having diamond jewelry, you may be unhappy to know that the most stable state of carbon under standard conditions is graphite. So as you read this, your diamond is turning to graphite, but fortunately the time scale for this conversion is many hundreds of years. If graphite is more stable, why were diamonds formed? The answer is that the formation of diamonds did not occur under standard conditions: it is well known that at very high temperatures and pressures, graphite can be converted to diamond. In the biological realm, one of the most critical reactions is the hydrolysis of ATP to ADP and P_i . Thermodynamics tells us that the equilibrium lies far toward ADP and P_i . However, solutions of ATP are quite stable under physiological conditions in the absence of the enzyme ATPase. A small amount of this enzyme will cause rapid and almost complete hydrolysis. Virtually all metabolic reactions occur much too slowly to sustain life in the absence of enzymes. Enzymes serve as catalysts and cause reactions to occur many orders of magnitude faster. Studies of the rates of hydrolysis under varying conditions allow us to say by what mechanism the reaction may occur, for example, how the substrates and products interact with the enzyme. Understanding the mechanism of biological processes is a research area of great current interest.

In this chapter, we will be interested in understanding some of the basic principles of chemical kinetics, with a few examples to illustrate the power of chemical kinetics. We will not delve deeply into the complex mathematics that is sometimes necessary nor the many specialized methods that are sometimes used to analyze chemical kinetics. Many treatises are available that provide a more complete discussion of chemical kinetics (1–5). Because the examples used will be relatively simple, they will not necessarily involve biological processes. The principles discussed, however, are generally applicable to all systems. The background presented here should be sufficient to get you started on utilizing chemical kinetics and reading the literature with some comprehension. The key three concepts that will be discussed first are *rates of chemical reactions*, *elementary reactions*, and *mechanisms of chemical reactions*.

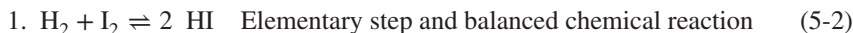
As a simple illustration, consider the reaction of hydrogen and iodine to give hydrogen iodide in the gas phase:



A possible mechanism for this reaction is for hydrogen and iodine to collide to produce hydrogen iodide directly. Another possible mechanism is for molecular iodine to first dissociate into iodine atoms and for the iodine atoms to react with hydrogen to produce hydrogen iodide. These possibilities can be depicted as

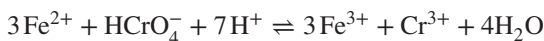


These two possible modes of reaction are quite distinct in terms of the chemistry occurring and are examples of two different mechanisms. We can also write these mechanisms in a more conventional manner:



As indicated, the individual steps in the mechanism are called *elementary steps*, and they must always add up to give the balanced chemical reaction. In the first mechanism, the elementary reaction and balanced chemical reaction happen to be the same, but they are conceptually quite distinct, as we shall amplify later.

A very important point to remember is that a balanced chemical reaction gives no information about the reaction mechanism. This is not so obvious in the aforementioned example, but consider the reaction



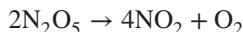
This clearly cannot be the reaction mechanism, as it would require the 11 reactants to encounter each other simultaneously, a very unlikely event.

5.2 REACTION RATES

The rate of a chemical reaction is a measure of how fast the concentration changes. If the concentration changes an amount Δc in time interval Δt , the rate of the reaction is $\Delta c/\Delta t$. If the limit of smaller and smaller time intervals is taken, this becomes dc/dt . If we apply this definition to Eq. (5-1), starting with H_2 and I_2 as reactants, the rate could be written as $-d[\text{H}_2]/dt$, $-d[\text{I}_2]/dt$, or $+d[\text{HI}]/dt$. For every mole of hydrogen and iodine consumed, 2 mol of hydrogen iodide are formed, so that this definition of the rate does not provide a unique definition: the value of the rate depends on which reaction component is under consideration. The rate of appearance of HI is twice as great as the rate of disappearance of H_2 or I_2 . This ambiguity is not convenient, so a convention is used to define the reaction rate, namely, the rate of change of the concentration divided by its coefficient in the balanced chemical reaction. This results in a unique reaction rate, R , for a given chemical reaction. By convention, R is always positive. For the case under consideration,

$$R = -\frac{d[\text{H}_2]}{dt} = -\frac{d[\text{I}_2]}{dt} = \frac{1}{2} \frac{d[\text{HI}]}{dt}$$

Consider a more complex chemical equation:



In this case, the reaction rate is

$$R = -\frac{1}{2} \frac{d[\text{N}_2\text{O}_5]}{dt} = \frac{1}{4} \frac{d[\text{NO}_2]}{dt} = \frac{d[\text{O}_2]}{dt}$$

Measuring the concentration as it changes with time can be very difficult, and some of our greatest advances in understanding chemical reactions have resulted from the development of new techniques for measuring the rates of chemical reactions, particularly the rates of very fast reactions. The most convenient method of measuring

reaction rates is to mix the reactants together and to measure the subsequent change in concentrations continuously using a spectroscopic technique such as light absorption. It is sometimes not possible to find a physical property to monitor continuously, so it may be necessary to stop the reaction and measure the concentration chemically or through radioactive tracers. We will not dwell on this point, except to stress that measuring the rate of a chemical reaction may not be trivial.

Once a method has been established for determining the reaction rate, the next step is to measure the dependence of the reaction rate on the concentrations of the reactants. The dependence of the rate on the concentrations is called the *rate law*. *The rate law cannot be predicted from the balanced chemical equation. It must be determined experimentally.* A few examples will serve to illustrate this point. Consider the hydrolysis of ATP catalyzed by an ATPase. The overall reaction is



and the observed dependence of the rate on concentration of ATP and enzyme in some cases is

$$R = \frac{k[\text{E}_0]}{1 + \frac{K_m}{[\text{ATP}]}} \quad (5-5)$$

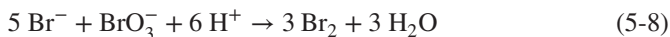
where $[\text{E}_0]$ is the total enzyme concentration and k and K_m are constants. The rate law certainly cannot be deduced from the overall chemical reaction. In fact, the rate law is not the same for all ATPases. Many different ATPases are found in biological systems, and they do not all hydrolyze ATP by the same mechanism. As another example, consider the simple reaction



The experimentally determined rate law is

$$R = \frac{k[\text{H}_2][\text{Br}_2]^{1/2}}{1 + \frac{K[\text{HBr}]}{[\text{Br}_2]}} \quad (5-7)$$

As a final example, consider the redox reaction



The observed rate law is

$$R = k[\text{Br}^-][\text{BrO}_3^-][\text{H}^+]^2 \quad (5-9)$$

These examples should emphasize the futility of attempting to predict the rate law from the balanced chemical equation.

The exponent of the concentration in the rate law is the *reaction order* with respect to that component (i.e., first order, second order, etc.). In some cases, such as Eq. (5-9), the concept of reaction order has a simple meaning. The rate law is first order with

respect to Br^- and BrO_3^- , and second order with respect to H^+ . For more complex rate laws, such as Eqs. (5-5) and (5-7), the concept of reaction order cannot be used for all of the concentrations, except in limiting conditions. For Eq. (5-7), the rate law is 1/2 order with respect to Br_2 at high concentrations of Br_2 and/or low concentrations of HBr . The lowercase k 's in the rate laws are called *rate constants*. The dimensions of rate constants can be deduced from the rate law by remembering that the rate is usually measured as M/s . In Eq. (5-9), k , therefore, has the dimensions of $\text{M}^{-3}\text{s}^{-1}$. The capital K 's in the aforementioned equations also are constants, but they are combinations of rate constants rather than individual rate constants.

5.3 DETERMINATION OF RATE LAWS

Many different methods exist for determining rate laws. Only a few methods are considered here. With the routine use of computers, numerical integration of the differential equations and simultaneous fitting of the data are possible. However, as with any experimental approach, it is best to first fit the data to simple models before embarking on complex computer fitting. Computer programs will always fit the data, but it is important to be sure that the fit is a good one and the proposed rate law makes sense.

Probably, the simplest and still often used method is the determination of initial rates. With this method, the rate of the reaction is measured at the very beginning of the reaction under conditions where the decrease in the concentrations of the reactants is so small that their concentrations in the rate law can be assumed to be the starting concentrations. In practice, this means that the concentrations should not change by more than a few percent. To illustrate the method, let us assume the rate law is

$$R = k(c_1)^a(c_2)^b \quad (5-10)$$

If c_2 is held constant and the initial rate is measured for different concentrations of c_1 , the coefficient a can be determined. For example, if the concentration c_1 is doubled and the rate increases by a factor of 4, a must be equal to 2. The same type of experiment can be carried out to determine b , namely, c_1 is held constant and c_2 is varied. The rate constant can readily be calculated, once the coefficients a and b are known, from the relationship $k = R/[(c_1)^a(c_2)^b]$. The determination of initial rates is especially useful for studying enzymatic reactions as we shall see later.

One of the most useful methods for determining the rate law is integration of the rate equation to obtain an analytical expression for the time dependence of the concentrations. The analytical equation is then compared with experimental data to see if it accurately describes the time dependence of the concentrations. As mentioned previously, computers can perform this integration numerically. We will consider two examples of integrated rate equations to illustrate the method.

The decomposition of nitrogen pentoxide can be written as



Assume that the decomposition is a first-order reaction:

$$R = -\frac{d[\text{N}_2\text{O}_5]}{dt} = k[\text{N}_2\text{O}_5] \quad (5-12)$$

This equation can easily be integrated

$$-\frac{d[\text{N}_2\text{O}_5]}{[\text{N}_2\text{O}_5]} = kdt$$

or

$$-\int_{[\text{N}_2\text{O}_5]_0}^{[\text{N}_2\text{O}_5]} \frac{d[\text{N}_2\text{O}_5]}{[\text{N}_2\text{O}_5]} = \int_0^t kdt$$

and

$$\ln\left(\frac{[\text{N}_2\text{O}_5]}{[\text{N}_2\text{O}_5]_0}\right) = -kt \quad (5-13)$$

In Eq. (5-13), $[\text{N}_2\text{O}_5]_0$ is the concentration when $t = 0$. As illustrated in Figure 5-1, this equation predicts that a plot of $\ln[\text{N}_2\text{O}_5]$ versus t should be a straight line with a slope of $-k$. The data, in fact, conform to this rate equation under most conditions.

This analysis can also be carried out for higher order reactions. Hydrogen iodide will react to give hydrogen and iodine under certain conditions:



The rate law is

$$R = -\frac{1}{2} \frac{d[\text{HI}]}{dt} = k[\text{HI}]^2 \quad (5-15)$$

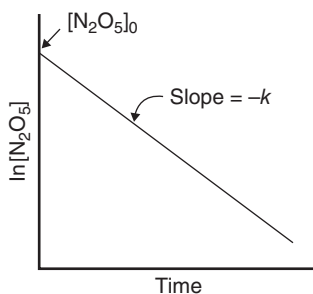


FIGURE 5-1. Demonstration of first-order kinetics. Plot of $\ln[\text{N}_2\text{O}_5]$ versus time according to Eq. (5-13). The straight line has a slope of $-k$ and an intercept on the y axis of $\ln[\text{N}_2\text{O}_5]_0$.

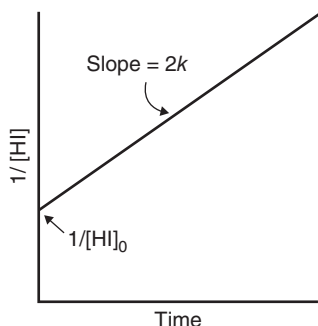


FIGURE 5-2. Demonstration of second-order kinetics. Plot of $1/[\text{HI}]$ versus time according to Eq. (5-16). The straight line has a slope of $2k$.

Rearrangement of this equation gives

$$-\frac{d[\text{HI}]}{[\text{HI}]^2} = 2k \, dt$$

which can be integrated to give

$$\frac{1}{[\text{HI}]} - \frac{1}{[\text{HI}]_0} = 2kt \quad (5-16)$$

Here $[\text{HI}]_0$ is the concentration when $t=0$. As illustrated in Figure 5-2, this equation predicts that a plot of $1/[\text{HI}]$ versus t should be a straight line with a slope of $2k$. In addition, the experimental data conform to this prediction.

It is frequently, but not always, possible to integrate rate equations analytically. Good experimental design can help to make the rate law relatively simple and therefore easy to integrate. In fact, experienced researchers will try to make the reaction first order whenever possible. This might seem like a major restriction but it is not. It is often possible to convert complex rate laws to conform to *pseudo-first-order kinetics*. For example, assume the rate law is

$$R = k(c_1)f(c_2)$$

where $f(c_2)$ is a function of the concentration c_2 . The function could be very complex or a simple power of the concentration—it does not matter. If the concentration c_2 is made much larger than the concentration c_1 , then it can be assumed to remain constant throughout the reaction, and the rate law becomes

$$R = k'(c_1)$$

where $k' = kf(c_2)$. The “constant” $kf(c_2)$ is called a “pseudo” first-order rate constant since it is constant under the experimental conditions used, but actually depends on

the concentration c_2 . If the concentration of c_2 is varied, the dependence of $f(c_2)$ on c_2 can be determined. It might seem restrictive to have a high concentration of c_2 , but in practice it is often sufficient to have the concentration of c_2 only about a factor of 10 higher than c_1 . Careful experimental design can make the job of determining the rate law much easier!

In determining the rate law, it is often necessary to use a broad range of experimental conditions before trying to interpret the rate law in terms of a chemical mechanism. As an illustration, consider the reaction



At constant pH, the rate law determined experimentally is

$$R = k[\text{I}^-][\text{OCl}^-] \quad (5-18)$$

If the pH is varied, it is found that k also varies. It was determined that $k = k'/[\text{OH}^-]$, where k' is a constant. Thus, a more complete rate law is

$$R = \frac{k'[\text{I}^-][\text{OCl}^-]}{[\text{OH}^-]} \quad (5-19)$$

The more information that can be determined about the concentration dependence of the rate the better the mechanism that can be postulated.

5.4 RADIOACTIVE DECAY

A good example of a first-order rate process is radioactive decay. Radioactive isotopes are frequently used in biological research. For example, the radioactive isotope of phosphorus, ^{32}P , gives off radiation according to the reaction



where β^- is a high-energy electron. The rate law for radioactive decay is

$$-\frac{d[^{32}\text{P}]}{dt} = k[^{32}\text{P}] \quad (5-21)$$

Integration of this rate law as performed previously (Eq. (5-13)) gives

$$[^{32}\text{P}] = [^{32}\text{P}]_0 e^{-kt} \quad (5-22)$$

or

$$\ln \left(\frac{[^{32}\text{P}]}{[^{32}\text{P}]_0} \right) = -kt$$

The rate of radioactive decay is usually given in terms of the half-life of the radioactive decay; in this case, the half-life is 14.3 days. In terms of the integrated rate law, Eq. (5-22), when half of the original radioactivity has decayed, $\ln(1/2) = -kt$, or the half-life for decay is $t_{1/2} = (\ln 2)/k$.

Thus, for radioactive decay, the half-life is constant. It does not matter when we start counting or how much radioactivity we start with, the radioactivity will decay to half of its original value in 14.3 days. This is a special property of first-order rate processes and is not true for other reaction orders where the half-life depends on the concentrations of the reactants. This is very convenient because regardless of when we start observing the rate of reaction, at $t = 0$, or at $t =$ any value, the integrated rate law is a simple single exponential.

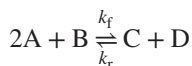
5.5 REACTION MECHANISMS

In general, many mechanisms are possible for a given reaction. Mechanisms are proposals for how the reaction occurs. A proposed mechanism must be consistent with the experimentally observed rate law, but this is usually true for many mechanisms. *Kinetic studies can disprove a mechanism but cannot prove a mechanism.* As a practical matter, the simplest mechanism consistent with all of the data is most appropriate, but at the end of the day, all that can be said is that a specific mechanism is consistent with known data. It is not possible to say that this *must* be “the” mechanism.

A mechanism consists of a combination of *elementary steps*, which must sum up to give the overall reaction. For an elementary step, the order and molecularity, the number of molecules involved in the reaction, are the same. Therefore, for elementary steps, the rate law can be written as the product of the concentrations of all reactants, each raised to the power of their stoichiometric coefficient, multiplied by a rate constant. Some examples of elementary steps and associated rate laws are given as follows:



Remember, rate laws can be derived from the chemical equation *only* for elementary steps and never for the balanced chemical equation of the overall reaction. If an elementary reaction is reversible, then the rate law is the difference between the rates of the forward and reverse reactions. Therefore, for the elementary step



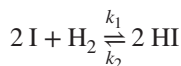
$$R = k_f[\text{A}]^2[\text{B}] - k_r[\text{C}][\text{D}]$$

We now have two criteria for a possible mechanism: (1) it must be consistent with the observed rate law and (2) the elementary steps must add up to give the overall

balanced chemical reaction. Let us return to the two proposed mechanisms for the reaction of H_2 and I_2 , Eqs. (5-2) and (5-3). The first mechanism contains only a single elementary step so that the rate law is

$$R = k_f[\text{H}_2][\text{I}_2] - k_r[\text{HI}]^2 \quad (5-23)$$

The second mechanism contains two elementary steps so that some assumptions must be made to derive the rate law. We will assume that iodine atoms are in rapid equilibrium with molecular iodine or to be more specific that this equilibrium is adjusted much more rapidly than the reaction of iodine atoms with molecular hydrogen. Furthermore, the concentration of iodine atoms is assumed to be much less than that of molecular iodine. These assumptions are, in fact, known to be correct. If we now consider the second elementary step in the mechanism,



the rate law can be written as

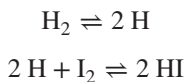
$$R = k_1[\text{I}]^2[\text{H}_2] - k_2[\text{HI}]^2 \quad (5-24)$$

but since the first step is always at equilibrium throughout the course of the reaction, $[\text{I}]^2 = K[\text{I}_2]$, where K is the equilibrium constant for the dissociation reaction. Substituting this relationship into Eq. (5-24) gives

$$R = k_1 K [\text{I}_2][\text{H}_2] - k_2 [\text{HI}]^2 \quad (5-25)$$

Equations 5-23 and 5-25 are identical in form as only the definitions of the constants are different, $k_f = k_1 K$ and $k_r = k_2$. Thus, we have shown that both mechanisms are consistent with the experimentally determined rate law, and therefore both are possible mechanisms. Even for this simple reaction, there remains a debate as to which is the more likely mechanism.

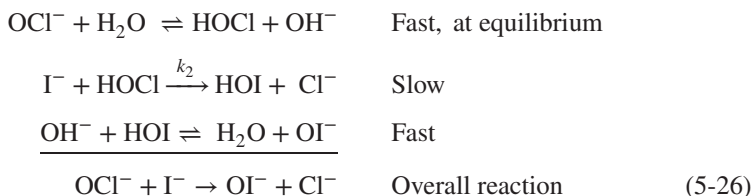
It is easy to postulate a third possible mechanism with the following elementary steps:



By analogy with the mechanism involving iodine atoms, it can be seen that this mechanism would give the rate law of Eq. (5-25) with K now being the dissociation constant for molecular hydrogen in equilibrium with hydrogen atoms. However, K is a known constant and if this constant is combined with the value of $k_1 K$ determined experimentally, k_1 can be calculated. It is found that the value of k_1 is much larger than the rate constant characterizing the maximum rate at which two hydrogen atoms and molecular iodine encounter each other in the gas phase. Therefore, this

mechanism can be ruled out as inconsistent with well-established theory. Thus, we have a third criterion for disproving a mechanism.

As a second example of how mechanisms can be deduced, let us return to the reaction of I^- and OCl^- to produce OI^- and Cl^- , Eq. (5-17). The experimentally determined rate law is given by Eq. (5-19). What is a possible mechanism? One possibility is the following scheme:



This mechanism is consistent with the balanced chemical equation. Now we must show that it is consistent with the observed rate law. The rate of a reaction is determined by the slowest step in the mechanism. Therefore, the rate of the overall reaction is given by the rate of the second elementary step:

$$R = k_2[\text{I}^-][\text{HOCl}] \quad (5-27)$$

If the first step is assumed to be fast and the concentration of HOCl small relative to that of I^- and Cl^- , then the concentration of HOCl can be calculated from the equilibrium relationship

$$\begin{aligned}
 K &= \frac{[\text{HOCl}][\text{OH}^-]}{[\text{OCl}^-]} \\
 [\text{HOCl}] &= \frac{K[\text{OCl}^-]}{[\text{OH}^-]}
 \end{aligned}$$

Insertion of this relationship into Eq. (5-27) gives

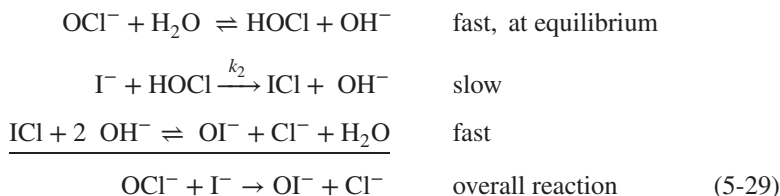
$$R = \frac{k_2 K [\text{I}^-][\text{OCl}^-]}{[\text{OH}^-]} \quad (5-28)$$

which is the observed rate law. Since this mechanism is consistent with the observed rate law and the balanced chemical equation, it is a possible mechanism.

You may very well be wondering if the creation of a mechanism is black magic. It is true that imagination and knowledge are important factors, but logic can be used. A species in the denominator results from a fast equilibrium prior to a rate-determining step. Therefore, what needs to be found is a first step involving one of the reactants that produces the desired species, in this case OH^- . The other product of the first step must then react with the second reactant. The remainder of the steps are fast reactions that are necessary to produce a balanced chemical reaction. Note that none of the steps after the rate-determining step play a role in determining the rate law. This is one of the simplest types of mechanisms, fast equilibria prior to and after a

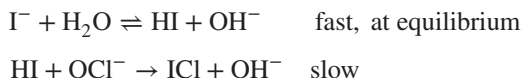
rate-determining step. Nature is not always so obliging, and more complex mechanisms in which several steps occur at comparable rates are often necessary to account for experimental findings.

Here is another mechanism, quite similar in concept:



Obviously, this gives the same rate law as the mechanism in Eq. (5-26), as all events prior to the rate-determining step are the same. However, the chemistry is quite different. In the first case, iodide attacks the oxygen; in the second case, it attacks the chlorine. These two mechanisms cannot be distinguished by kinetics. Both are equally consistent with the data.

Thus far, we have considered mechanisms in which OCl^- is the initial reactant. Now let us look at mechanisms in which I^- is the initial reactant such as



By analogy, it should be clear that this mechanism would give the correct rate law, and rapid reactions after the rate-determining step can be added to give the correct balanced chemical equation. However, the equilibrium constant for the first step is known, and if this is combined with the results of the kinetic experiments, the rate constant for the rate-determining step would exceed the theoretically possible value. This is because the concentration of HI is much, much smaller than that of HOCl. Thus, mechanisms of this type can be excluded on the basis of theoretical concepts.

These two examples illustrate how mechanisms can be related to the results of kinetic experiments. It is a great challenge to devise mechanisms. Once a mechanism is postulated, it is the job of the kineticist to devise experiments that will test the mechanism, often disproving it and requiring postulation of a new mechanism. However, it is important to remember that no matter how convincing the arguments, a mechanism cannot be proved. Thus far, we have considered very simple reactions that are not biological, as biological reactions are typically very complex. The purpose of this discussion has been to illustrate the principles and concepts of chemical kinetics. We will later consider the kinetic analysis of enzymatic reactions, a very relevant and timely subject.

5.6 TEMPERATURE DEPENDENCE OF RATE CONSTANTS

Before discussing the kinetic analyses of biological reactions, a few additional concepts will be described. Reaction rates are often dependent on the temperature, and

typically reactions go faster as the temperature increases. The first quantitative treatment of the temperature dependence of reaction rates was developed by Arrhenius in the late 1800s. He proposed that the temperature dependence of the rate constant could be described by the equation

$$k = A \exp \left(-\frac{E_a}{RT} \right) \quad (5-30)$$

where A is a constant, E_a is the *activation energy*, R is the gas constant, and T is the Kelvin temperature. This equation predicts that a plot of $\ln k$ versus $1/T$ should be a straight line with a slope of $-E_a/R$. This behavior is, in fact, followed in most cases. Equation 5-30 can be differentiated with respect to temperature to give

$$\frac{d(\ln k)}{dT} = \frac{E_a}{RT^2} \quad (5-31)$$

This equation can be integrated to give

$$\ln \left(\frac{k_{T_2}}{k_{T_1}} \right) = \frac{\left(\frac{E_a}{R} \right) (T_2 - T_1)}{T_1 T_2} \quad (5-32)$$

Equation 5-32 permits the rate constant to be calculated at any temperature if it is known at one temperature and the activation energy has been determined. Note that these equations are similar in form to those describing the temperature dependence of the equilibrium constant except that the activation energy has replaced the standard enthalpy of reaction. In some cases, the activation energy may change with temperature, thereby making the analysis more complex.

The physical model behind the Arrhenius equation is that an energy barrier, E_a , must be overcome in order for the reaction to occur. This is shown conceptually in Figure 5-3. The reaction “path” or “coordinate” can be envisaged as the approach of the reactants to each other that results in the lowest activation energy. The energy difference between products, P , and reactants, R , is ΔE . For simple reactions in the gas phase, the energy can be calculated as a function of the distance between reactants. These calculations define an energy surface, and the dynamic course of the reaction on this energy surface can be determined. Even in the gas phase, this can only be done for very simple reactions. For reactions in solution, this can only be considered a conceptual model.

You might guess from your knowledge of thermodynamics that the energy is probably not the best parameter to use to characterize the dynamics of a reaction. A theory has been developed that instead discusses the reaction path in terms of the Gibbs energy. This theory is called *transition state theory*. The basic postulate is that a *transition state* is formed in the course of the reaction, which is in equilibrium with the reactants. This model is shown schematically in Figure 5-4. In transition state theory, the reactants go through a transition state that is at a higher Gibbs energy than the reactants. They must go over this Gibbs energy barrier to produce products. In

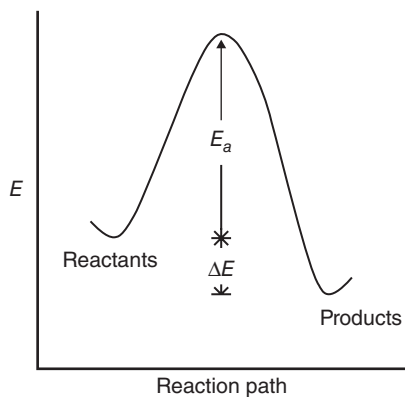


FIGURE 5-3. Schematic representation of the energy, E , versus the “reaction path” during the course of a reaction as discussed in the text. The definition of the activation energy, E_a , is indicated.

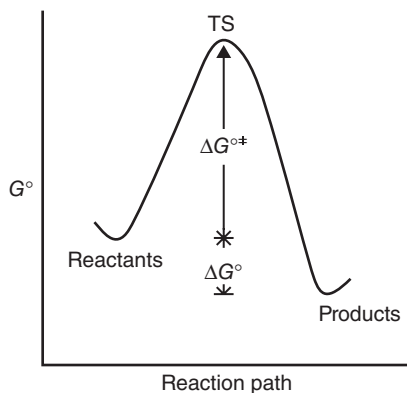


FIGURE 5-4. Schematic representation of the standard Gibbs energy, G° , versus the “reaction path” during the course of a reaction as discussed in the text. The Gibbs energy of activation, $\Delta G^{\circ\ddagger}$, is defined in this diagram. The transition state at the maximum of the Gibbs energy is indicated by TS. The reactants and products are at minima in the Gibbs energy curve.

Figure 5-4, the difference in Gibbs energy between products and reactants is ΔG° , consistent with thermodynamic principles. The Gibbs energy difference between the transition state and reactants, $\Delta G^{\circ\ddagger}$, is called the standard Gibbs energy of activation. In terms of transition state theory, the rate constant can be written as

$$k = \left(\frac{k_B T}{h} \right) \exp \left(-\frac{\Delta G^{\circ\ddagger}}{RT} \right) \quad (5-33)$$

where k_B is Boltzmann's constant and h is Planck's constant. (A more exact derivation includes an additional parameter, the transmission coefficient, which is often approximately equal to 1.) Since $\Delta G^{o\ddagger} = \Delta H^{o\ddagger} - T\Delta S^{o\ddagger}$, the rate constant can be restated as

$$k = \left(\frac{k_B T}{h} \right) \exp \left(\frac{\Delta S^{o\ddagger}}{R} \right) \exp \left(\frac{-\Delta H^{o\ddagger}}{RT} \right) \quad (5-34)$$

Here $\Delta S^{o\ddagger}$ is the standard entropy of activation and $\Delta H^{o\ddagger}$ is the standard enthalpy of activation.

If the standard entropy and enthalpy of activation are assumed to be temperature independent, Eq. (5-34) can be differentiated to give

$$\frac{d(\ln k)}{dT} = \frac{(\Delta H^{o\ddagger} + RT)}{RT^2} \quad (5-35)$$

Thus, transition state theory predicts a temperature dependence of the rate constant very similar to the Arrhenius theory with $E_a = \Delta H^{o\ddagger} + RT$. Since RT is usually small relative to the standard enthalpy of activation, the activation energy and standard enthalpy of activation are usually quite similar. The Arrhenius and transition state formulations cannot be differentiated by this small difference in the temperature dependence since both the activation energy and the standard enthalpy of activation can be temperature dependent. From our knowledge of thermodynamics, we know that the enthalpy of activation is temperature dependent if a heat capacity difference exists between the transition state and the reactants. If the temperature dependence of the rate constant is to be analyzed in terms of transition state theory, it is more convenient to plot $\ln(k/T)$ versus $1/T$ as $d \ln(k/T)/dT = \Delta H^{o\ddagger}/(RT^2)$ or $d \ln(k/T)/d(1/T) = -\Delta H^{o\ddagger}/R$.

A simple derivation of Eq. (5-33) is possible. The concentration of the transition state, TS, can be calculated from the relationship

$$\frac{[\text{TS}]}{[\text{Reactants}]} = \exp \left(\frac{-\Delta G^{o\ddagger}}{RT} \right)$$

where $[\text{Reactants}]$ represents the concentrations of the reactants raised to the appropriate powers for the stoichiometric equation of the elementary step. The rate of the reaction is the concentration of the transition state multiplied by the frequency with which it passes over the Gibbs energy barrier. This frequency of passage over the barrier can be derived from an analysis of the reaction coordinate with statistical mechanics and is given by $k_B T/h$. Therefore, the rate of reaction is

$$R = \left(\frac{k_B T}{h} \right) [\text{TS}] = [\text{Reactants}] \left(\frac{k_B T}{h} \right) \exp \left(\frac{-\Delta G^{o\ddagger}}{RT} \right)$$

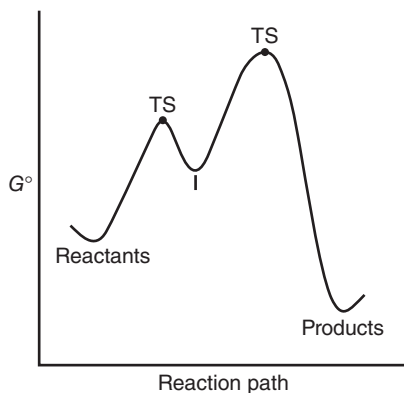


FIGURE 5-5. Schematic representation of the standard Gibbs energy, G° , versus the “reaction path” during the course of a two-step reaction. The intermediate, I, is at a minimum in the Gibbs energy whereas the transition state, TS, for each step is at a maximum. The reactants and products are also at minima in the free energy curve.

The rate constant in Eq. (5-33) follows directly from the aforementioned equation. This nonrigorous derivation provides a conceptual framework for the theory.

If a reaction goes through an intermediate, the intermediate would correspond to a minimum in the Gibbs energy, and each elementary step would have its own transition state. This is shown schematically in Figure 5-5 for a sequence of two elementary steps. The step with the highest Gibbs energy barrier is the rate-determining step.

Transition state theory is a very useful method for correlating and understanding kinetic studies. Because the framework of the theory is similar to thermodynamics, this produces a consistent way of discussing chemical reactions. The entropy and enthalpy of activation are often discussed in molecular terms. It should be remembered that, for kinetics as with thermodynamics, such interpretations should be approached with extreme caution.

5.7 RELATIONSHIP BETWEEN THERMODYNAMICS AND KINETICS

Obviously, the principles of thermodynamics and kinetics must be self-consistent. In fact, this places some useful restrictions on the relationships between rate constants. In order to see how rate constants are related to equilibrium constants, consider the elementary step



The rate of this reaction is

$$R = -\frac{d[A]}{dt} = k_1[A][B] - k_2[AB] \quad (5-37)$$

At equilibrium, the net rate of reaction must be zero. If $R = 0$, $k_1[A]_e[B]_e = k_2[AB]_e$, where the subscript e designates the equilibrium concentration. Thus, we see that

$$K = \frac{k_1}{k_2} = \frac{[AB]_e}{[A]_e[B]_e} \quad (5-38)$$

In this case, the equilibrium constant, K , is equal to the ratio of rate constants. Similar relationships between the rate constants and equilibrium constants can be found for more complex situations by setting the net rate equal to zero at equilibrium.

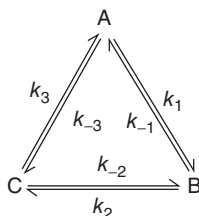
If Eq. (5-38) is cast into the framework of transition state theory, we obtain

$$K = \exp \left[- \frac{(\Delta G_1^{\ddagger} - \Delta G_2^{\ddagger})}{RT} \right] = \exp \left(- \frac{\Delta G^0}{RT} \right) \quad (5-39)$$

This result indicates the relationship between the standard Gibbs energy changes of activation and the standard Gibbs energy change for the reaction. This relationship can also be seen in the diagram of the Gibbs energy versus reaction path (Fig. 5-4).

A more subtle relationship can be found if reaction cycles occur because of the *principle of detailed balance* or *microscopic reversibility*. This principle states that a mechanism for the reaction in the forward direction must also be a mechanism for the reaction in the reverse direction. Furthermore, at equilibrium, the forward and reverse rates are equal along each reaction pathway. This means that once we have found a possible mechanism for the reaction in one direction, we have also found a possible mechanism for the reaction in the other direction.

To illustrate this principle, consider the following triangular reaction mechanism:



According to the principle of detailed balance, each of the individual reactions must be at equilibrium when equilibrium is attained, or

$$k_1[A]_e = k_{-1}[B]_e$$

$$k_2[B]_e = k_{-2}[C]_e$$

$$k_3[C]_e = k_{-3}[A]_e$$

If the right-hand sides of these equations are multiplied together and set equal to the left-hand sides of the equations multiplied together, we obtain

$$k_1 k_2 k_3 = k_{-1} k_{-2} k_{-3} \quad (5-40)$$

Thus, we find that the six rate constants are not independent, nor are the three equilibrium constants! This result may seem obvious, but many people have violated the principle of detailed balance in the literature. It is important to confirm that this principle is obeyed whenever reaction cycles are present.

5.8 REACTION RATES NEAR EQUILIBRIUM

The kinetics of reactions close to equilibrium merit special consideration. As we shall see, the rate laws become quite simple near equilibrium, and methods exist that permit very fast reactions to be studied near equilibrium. We will begin with an elementary step that is a reversible first-order reaction, such as protein denaturation in the middle of the transition from the native to the denatured state:



The rate equation characterizing this system is

$$-\frac{d[N]}{dt} = k_f[N] - k_r[D] \quad (5-42)$$

We will now introduce new concentration variables; that is, we will set each concentration equal to its equilibrium value plus the deviation from equilibrium:

$$[N] = [N]_e + \Delta[N]$$

$$[D] = [D]_e + \Delta[D]$$

Note that the equilibrium concentrations are constants, independent of time, and mass conservation requires that $\Delta[N] = -\Delta[D]$. Inserting these relationships into Eq. (5-42) gives

$$\begin{aligned} -\frac{d[\Delta N]}{dt} &= k_f([N]_e + \Delta[N]) - k_r([D]_e - \Delta[N]) \\ &= k_f[N]_e - k_r[D]_e + (k_f + k_r)\Delta[N] \\ &= (k_f + k_r)\Delta[N] = \Delta[N]/\tau \end{aligned}$$

In deriving the aforementioned relationship, use has been made of the relationship $k_f[N]_e = k_r[D]_e$ and $1/\tau = k_f + k_r$; τ is called the relaxation time. This first-order differential equation can be integrated as before to give

$$\Delta[N] = \Delta[N]_0 e^{-t/\tau} \quad (5-43)$$

where $\Delta[N]_0$ is the deviation of N from its equilibrium value at $t=0$.

Special note should be made of two points. First, the relaxation time can readily be obtained from experimental data simply by plotting $\ln(\Delta[N])$ versus t , and second, the

first-order rate constant characterizing this reaction, that is, the reciprocal relaxation time, is the sum of the two rate constants. If the equilibrium constant is known, both rate constants can be obtained from a single experiment.

How might such an experiment be carried out? In the case of protein denaturation, a small amount of denaturant such as urea might be added quickly to the solution. The ratio of the native and denatured proteins would then move to a new equilibrium value characteristic of the higher urea concentration. Alternatively, if thermal denaturation is being studied, the temperature might rapidly be raised, establishing a new equilibrium ratio. The rate of conversion of the native state to the denatured state can be measured after the experimental conditions are changed. For a reversible first-order reaction such as Eq. (5-41), the time course of the entire reaction follows a single exponential, and the effective rate constant is the sum of the two rate constants. A wide range of methods exist for changing the equilibrium conditions: concentration jumps and temperature jumps are particularly useful, but pressure jumps and electrical field jumps have also been used.

Studying rates near equilibrium is especially advantageous for higher order reactions and complex mechanisms. As a final example, we will consider the elementary step of an enzyme, E, combining with a substrate, S.



The rate equation characterizing this system is

$$-\frac{d[E]}{dt} = k_f[E][S] - k_r[ES] \quad (5-45)$$

We will write the concentrations as the sum of the equilibrium concentration plus the deviation from equilibrium.

$$[E] = [E]_e + \Delta[E]$$

$$[S] = [S]_e + \Delta[S]$$

$$[ES] = [ES]_e + \Delta[ES]$$

Furthermore, mass conservation requires that $\Delta[E] = \Delta[S] = -\Delta[ES]$. Inserting these relationships into Eq. (5-45) gives

$$-\frac{d(\Delta[E])}{dt} = \{k_f([E]_e + [S]_e) + k_r\}\Delta[E] + k_f[E]_e[S]_e - k_r[ES]_e + k_f(\Delta[E])^2$$

Near equilibrium, the deviation of concentrations from their equilibrium values is so small that the term $(\Delta[E])^2$ can be neglected—this simplification in the rate equation results from being near equilibrium. For example, if the deviation is 10% (0.1), the square of the deviation is 1% (0.01). Furthermore, $k_f[E]_e[S]_e = k_r[ES]_e$, so that the final rate equation is

$$-\frac{d\Delta[E]}{dt} = \frac{\Delta[E]}{\tau} \quad (5-46)$$

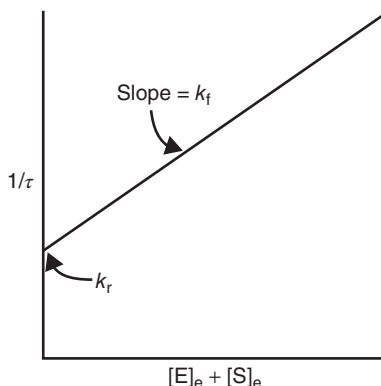


FIGURE 5-6. Schematic representation of a plot of the reciprocal relaxation time, $1/\tau$, versus the sum of the equilibrium concentrations, $[E]_e + [S]_e$, according to Eq. (5-47). As indicated, both of the rate constants can be obtained from the data.

with

$$\frac{1}{\tau} = k_f([E]_e + [S]_e) + k_r \quad (5-47)$$

Although the reaction of enzyme with substrate is a second-order reaction, the rate equation near equilibrium, Eq. (5-46), is the same as for a first-order reaction. In fact, all rate equations become first order near equilibrium! This is because only terms containing Δc are retained in the rate law: $(\Delta c)^2$ and higher powers are neglected. As before, Eq. (5-46) can easily be integrated, and the relaxation time can be obtained from the experimental data. If the relaxation time is determined at various equilibrium concentrations, a plot of $1/\tau$ versus $([E]_e + [S]_e)$ can be made, as shown schematically in Figure 5-6. The intercept is equal to k_r and the slope is equal to k_f . The study of reactions near equilibrium has been especially important for biological systems, as it has permitted the study of important elementary steps such as hydrogen bonding and proteolytic reactions, as well as more complex processes such as enzyme catalysis and ligand binding to macromolecules.

5.9 SINGLE MOLECULE KINETICS

Thus far we have considered the rates of reactions for systems containing a large number of molecules, i.e., approximately Avogadro's number. However, modern advances in light microscopy utilizing fluorescence detection now permit the kinetics of reaction of individual molecules to be studied (cf. 5). By the use of lasers, the reaction volume viewed can be reduced to microliters so that single molecules can be detected by observing their optical properties such as fluorescence. For conventional systems, the measured rate constant is an average of the rates of reaction of the individual molecules. If the rates of reaction of individual molecules are measured, they will vary over a wide range. For a simple first order reaction, say A going to B, the link

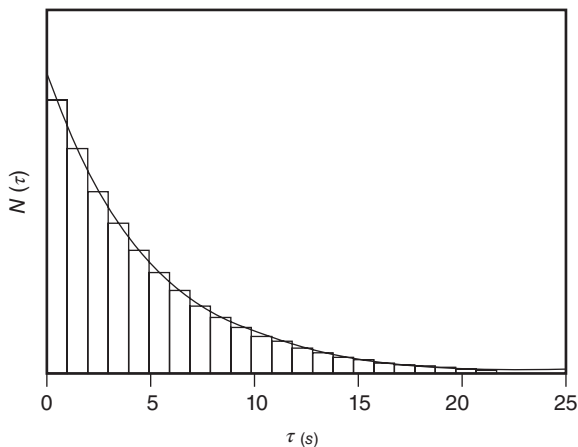


FIGURE 5-7. Histogram of the fraction of kinetic events with a lifetime τ versus τ . These are hypothetical data from a single molecule experiment. The line through the data has been calculated with Eq. (5-48) and $k = 0.2 \text{ s}^{-1}$.

between the system or ensemble rate constant and the individual reaction lifetimes is given by the relationship

$$\frac{N(\tau)}{N_{\text{Total}}} = k \exp(-k\tau) \quad (5-48)$$

Here $N(\tau)$ is the number of kinetic events with lifetime τ , N_{Total} is the total number of events (lifetimes) observed, and k is the ensemble rate constant. The experiments consist of measuring hundreds, sometimes thousands, of lifetimes of individual molecules, and then sorting them into ranges, e.g., the number of lifetimes in the range 0 to 1 s, 1–2 s, etc. The data are then plotted as a bar graph, and the midpoints of the bars are analyzed with Eq. (5-48). A plot of this type is shown in Figure 5-7. Thus the reaction times of individual molecules can be related to conventional rate constants.

What is the advantage of single molecule measurements over conventional kinetic studies? In some cases, reactions can be observed that are masked by the averaging process. For example, consider a molecule moving between or reacting with different sites on a molecule such as DNA. In ensemble experiments, many different reactions would be occurring at the same time, and sorting them all out would be difficult, if not impossible, because all of the different reactions are initiated at a different time. However, with single molecule experiments, the reactant molecule could be added to a single molecule of DNA to initiate the reaction and/or sequence of reactions, and the events on a single molecule can be observed as the reaction evolves. As a second example, consider an overall reaction at equilibrium with several different reactions involved in the equilibration. Single molecule measurements have the potential of distinguishing among these reactions. The time resolution of single molecule experiments is extremely broad, ranging from nanoseconds to many seconds.

The measurement of reaction rates with single molecule methods is a rapidly developing field of research and shows great promise for revealing details of reaction mechanisms that cannot be probed with conventional kinetics.

REFERENCES

1. I. Tinoco, Jr., K. Sauer, & J. C. Wang, J.D. Puglisi, G. Harbison and D. Rovnyak, *Physical Chemistry: Principles and Applications in Biological Sciences*, 5th edition, Prentice Hall, Englewood Cliffs, NJ, 2013.
2. C. R. Cantor and P. R. Schimmel, *Biophysical Chemistry*, Freeman, San Francisco, CA, 1980.
3. G. G. Hammes, *Enzyme Catalysis and Regulation*, Academic Press, New York, 1982.
4. P. L. Houston, *Chemical Kinetics and Reaction Dynamics*, Courier Corporation (Dover), 2006.
5. R. D. Smiley and G. G. Hammes, *Chem. Rev.* 45, 3080 (2006).

PROBLEMS

- 5-1.** The activity of the antibiotic penicillin decomposes slowly when stored in a buffer at pH 7.0, 298 K. The time dependence of the penicillin antibiotic activity is given in the following table.

Time (weeks)	Penicillin activity (arbitrary units)
0	10,100
1.00	8180
2.00	6900
3.00	5380
5.00	3870
8.00	2000
10.00	1330
12.00	898
15.00	403
20.00	167

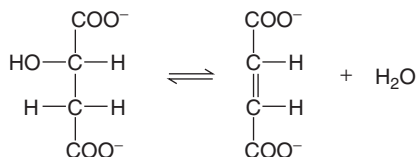
What is the rate law for this reaction, that is, what is the order of the reaction with respect to the penicillin concentration? Calculate the rate constant from the data if possible. (Data adapted from Ref. 1.)

- 5-2.** The kinetics of the reaction $2\text{Fe}^{3+} + \text{Sn}^{2+} \rightarrow 2\text{Fe}^{2+} + \text{Sn}^{4+}$ has been studied extensively in acidic aqueous solutions. When Fe^{2+} is added initially at relatively high concentrations, the rate law is

$$R = \frac{k[\text{Fe}^{3+}]^2[\text{Sn}^{2+}]}{[\text{Fe}^{2+}]}$$

Postulate a mechanism that is consistent with this rate law. Show that it is consistent by deriving the rate law from the proposed mechanism.

- 5-3. The conversion of L-malate to fumarate is catalyzed by the enzyme fumarase:



The nonenzymatic conversion is very slow in neutral and alkaline media and has the rate law

$$R = \frac{k[\text{L-malate}]}{[\text{H}^+]}$$

Postulate two mechanisms for the nonenzymatic conversion and show that they are consistent with the rate law.

- 5-4. The radioactive decay rates of naturally occurring radioactive elements can be used to determine the age of very old materials. For example, $^{14}_6\text{C}$ is radioactive and emits a low-energy electron with a half-life of about 5730 years. Through a balance of natural processes, the ratio of $^{14}\text{C}/^{12}\text{C}$ is constant in living organisms. However, in dead organisms or material, this ratio decreases as the ^{14}C decays. Since the radioactive decay is known to be a first-order reaction, the age of the material can be estimated by measuring the decrease in the $^{14}\text{C}/^{12}\text{C}$ ratio. Suppose a piece of ancient wool is found in which the ratio has been found to decrease by 20%. What is the age of the wool?

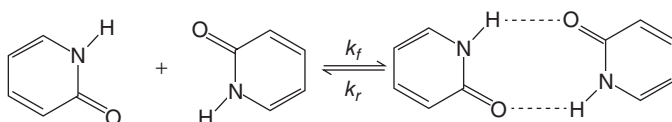
- 5-5. The nonenzymatic hydration of CO_2 can be written as



The reaction is found to be first order in both directions. Because the water concentration is constant, it does not appear in the expression for the equilibrium or rate equation. The first-order rate constant in the forward direction has a value of 0.0375 s^{-1} at 298 K and 0.0021 s^{-1} at 273 K. The thermodynamic parameters for the equilibrium constant at 298 K are $\Delta H^\circ = 1.13 \text{ kcal/mol}$ and $\Delta S^\circ = -8.00 \text{ cal/(mol K)}$.

- Calculate the Arrhenius activation energy for the rate constant of the forward reaction. Also calculate the enthalpy and entropy of activation according to transition state theory at 298 K.
- Calculate the rate constant for the reverse reaction at 273 and 298 K. Assume that ΔH° is independent of the temperature over this temperature range.
- Calculate the Arrhenius activation energy for the rate constant of the reverse reaction. Again, also calculate the enthalpy and entropy of activation at 298 K.

- 5-6.** A hydrogen bonded dimer is formed between 2-pyridone according to the reaction



The relaxation time for this reaction, which occurs in nanoseconds, has been determined in chloroform at 298 K at various concentrations of 2-pyridone. The data obtained are [G. G. Hammes and A. C. Park, *J. Am. Chem. Soc.* **91**, 956 (1969)]

2-Pyridone (M)	$10^9 \tau$ (s)
0.500	2.3
0.352	2.7
0.251	3.3
0.151	4.0
0.101	5.3

From these data, calculate the equilibrium and rate constants characterizing this reaction. *Hint:* If the expression for the relaxation time is squared, the concentration dependence can be expressed as a simple function of the total concentration of 2-pyridone.

Applications of Kinetics to Biological Systems

6.1 INTRODUCTION

We will now consider some applications of kinetic studies to biological systems. This discussion will center on enzymes, as kinetic analyses of enzymes represent a major research field and have provided considerable insight into how enzymes work. Enzymes are proteins that are incredibly efficient catalysts: they typically increase the rate of a chemical reaction by six orders of magnitude or more. Understanding how this catalytic efficiency is achieved and the exquisite specificity of enzymes has intrigued biologists for many decades and still provides a challenge for modern research. Since enzyme deficiencies are the source of many diseases, enzymes are also a target for medical research and modern therapeutics. The field of enzyme kinetics and mechanisms is so vast that we will present only an abbreviated discussion. More complete discussions are available (cf. Refs. 1–3).

We will also discuss catalysis by RNA (ribozymes), which is important in the processing of RNA in biological systems, as well as some kinetic studies of DNA denaturation and renaturation.

6.2 ENZYME CATALYSIS: THE MICHAELIS–MENTEN MECHANISM

We now consider the analysis of a simple enzymatic reaction. This discussion will introduce some new concepts that are particularly useful for analyzing complex systems. The conversion of a single substrate to product will be taken as a prototype reaction:



A typical example is the hydration of fumarate to give L-malate (and the reverse dehydration reaction) catalyzed by the enzyme fumarase. If a very low concentration of enzyme is used relative to the concentration of substrate, a plot of the initial rate of

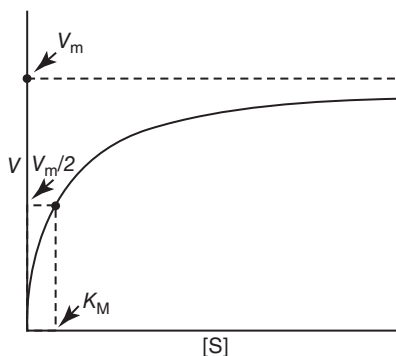
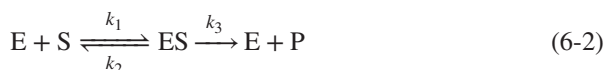


FIGURE 6-1. A plot of the initial velocity, v , versus the substrate concentration, $[S]$, for an enzymatic reaction that can be described by the Michaelis–Menten mechanism.

the reaction, or initial velocity, v , is hyperbolic, as shown in Figure 6-1. The limiting initial velocity at high concentrations of substrate is called the *maximum velocity*, V_m , and the concentration of substrate at which the initial velocity is equal to $V_m/2$ is called the *Michaelis constant*, K_M .

A mechanism that quantitatively accounts for the dependence of the initial velocity on substrate concentration was postulated by Michaelis and Menten. It can be written in terms of elementary steps as



where E is the enzyme and ES is a complex consisting of the enzyme and substrate. The total concentration of enzyme, $[E_0]$, is assumed to be much less than the total concentration of substrate, $[S_0]$. This mechanism is an example of *catalysis* in that the enzyme is not consumed in the overall reaction, which is greatly accelerated by the enzyme. If we tried to do an exact mathematical analysis of the mechanism in Eq. (6-2), coupled differential equations would need to be solved. Fortunately, this is not necessary. We will consider two approximations that can adequately account for the data.

The first is the *equilibrium approximation*. With this approximation, the first step is assumed to be always at equilibrium during the course of the reaction. This means that the equilibration of the first step is much more rapid than the breakdown of ES to P , or $k_3 \ll k_2$. We then have

$$\frac{d[P]}{dt} = v = k_3[ES] \quad (6-3)$$

and

$$\frac{k_1}{k_2} = \frac{[ES]}{[E][S]} \quad (6-4)$$

Furthermore, conservation of mass requires that

$$\begin{aligned} [E_0] &= [E] + [ES] \\ &= [ES] \left(1 + \frac{[E]}{[ES]} \right) = [ES] \left\{ 1 + \frac{k_2}{k_1 [S]} \right\} \end{aligned} \quad (6-5)$$

or

$$[ES] = \frac{[E_0]}{\left\{ 1 + \frac{k_2}{k_1 [S]} \right\}} \quad (6-6)$$

Substitution of Eq. (6-6) into (6-3) gives

$$\frac{d[P]}{dt} = v = \frac{k_3 [E_0]}{1 + \frac{k_2}{k_1 [S]}} \quad (6-7)$$

which can be rewritten as

$$v = \frac{V_m}{1 + K_M/[S]} \quad (6-8)$$

with

$$V_m = k_3 [E_0] \quad (6-9)$$

and

$$K_M = \frac{k_2}{k_1} \quad (6-10)$$

Equation (6-8) is found to account quantitatively for the data. The maximum velocity is proportional to the total enzyme concentration, and when $K_M = [S]$, $v = V_m/2$, as required.

In some cases, it is possible to obtain independent estimates or measurements of the equilibrium constant for the first step. Sometimes this independent measurement is in good agreement with the constant obtained from kinetic studies, sometimes it is not. Obviously, when the Michaelis constant and the equilibrium constant are not in agreement, the mechanism must be reexamined.

A more general analysis of the Michaelis–Menten mechanism makes use of the *steady-state approximation*. This approximation does not make any assumptions about the relative values of the rate constants but assumes that the concentrations of E and ES are very small relative to S, consistent with the experimental conditions. Under these conditions, it is assumed that the rate of change of the concentrations of E and ES is very small relative to the rate of change of the concentration of S, or

$$\frac{d[E]}{dt} = \frac{d[ES]}{dt} \approx 0$$

This is the steady-state approximation, an approximation that is very important for analyzing complex mechanisms. A careful mathematical analysis shows that this is a

very good approximation under the experimental conditions used, namely, when the total substrate concentration is much greater than the total enzyme concentration.

What does the steady-state approximation mean? If we go back to the mechanism (Eq. (6-2)), we find that

$$-\frac{d[\text{ES}]}{dt} = (k_2 + k_3)[\text{ES}] - k_1[\text{E}][\text{S}] \approx 0$$

or

$$\frac{k_1}{(k_2 + k_3)} = \frac{[\text{ES}]}{[\text{E}][\text{S}]} \quad (6-11)$$

Note the similarity of this equation to the equilibrium constant (Eq. (6-4)). This means that the ratio of concentrations remains constant. However, the ratio is not the equilibrium concentrations; it is the steady-state concentrations. If $k_3 \ll k_2$, the steady state is the same as the equilibrium state, so that the equilibrium condition is a special case of the steady state. For all other steady-state situations, the ratio of concentrations is always less than the equilibrium ratio.

We can now calculate the rate law for the steady-state approximation exactly as for the equilibrium approximation.

$$[\text{E}_0] = [\text{ES}] \left(1 + \frac{[\text{E}]}{[\text{ES}]} \right) = [\text{ES}] \left\{ 1 + \frac{(k_2 + k_3)}{k_1[\text{S}]} \right\} \quad (6-12)$$

$$v = \frac{d[\text{P}]}{dt} = k_3[\text{ES}] = \frac{k_3[\text{E}_0]}{1 + \frac{(k_2 + k_3)}{k_1[\text{S}]}}$$

$$v = \frac{V_m}{1 + K_M/[\text{S}]} \quad (6-13)$$

with

$$V_m = k_3[\text{E}_0] \quad (6-14)$$

and

$$K_M = \frac{(k_2 + k_3)}{k_1} \quad (6-15)$$

Thus, it is clear that the equilibrium approximation and steady-state approximation give rate laws that are indistinguishable experimentally. Most generally, the Michaelis constant is a steady-state constant, but in a limiting case it can be an equilibrium constant. As previously stated, both situations have been observed. The steady-state approximation is more general than the equilibrium approximation and is typically employed in the analysis of enzyme mechanisms.

With modern computers, data analysis is very easy, and experimental data can be fit directly to Eq. (6-8) (or Eq. (6-13)) by a nonlinear least squares fitting procedure. However, it is always a good idea to be sure that the data indeed do conform to the

equation of choice by a preliminary analysis. Equation (6-8) can be linearized by taking its reciprocal:

$$\frac{1}{v} = \frac{1}{V_m} + \frac{K_M}{V_m[S]} \quad (6-16)$$

This is called the Lineweaver-Burk equation. A plot of $1/v$ versus $1/[S]$ is linear, and the slope and intercept can be used to calculate the maximum velocity and Michaelis constant. This equation can be used for the final data analysis providing proper statistical weighting is used. An alternative linearization of Eq. (6-8) is to multiply Eq. (6-16) by $[S]$ to give

$$\frac{[S]}{v} = \frac{[S]}{V_m} + \frac{K_M}{V_m} \quad (6-17)$$

This equation predicts that a plot of $[S]/v$ versus $[S]$ should be linear and gives a better weighting of the data than the double reciprocal plot (Eq. (6-16)). Obviously, the final results should be independent of how the data are plotted and analyzed! Examples of both plots are given in Figure 6-2.

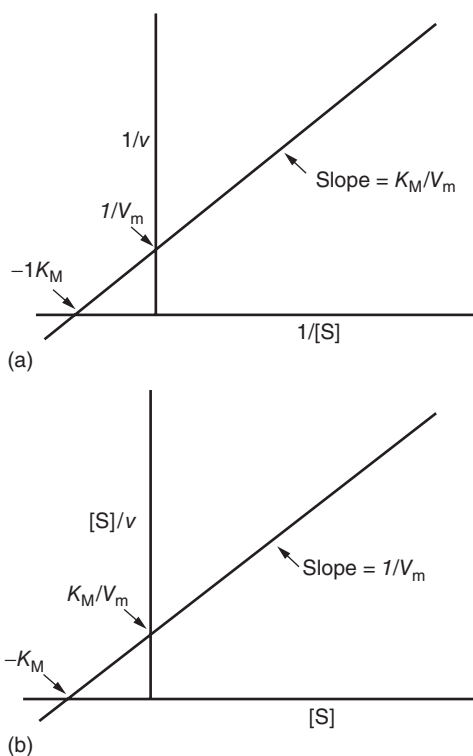


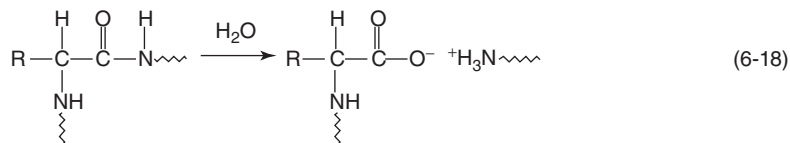
FIGURE 6-2. Linear plots commonly used for analyzing data that follow Michaelis-Menten kinetics. (a) Plot of the reciprocal initial velocity, v , versus the reciprocal substrate concentration, $1/[S]$. (b) Plot of $[S]/v$ versus $[S]$.

The steady-state analysis can be extended to much more complex enzyme mechanisms, as well as to other biological processes, but this is beyond the scope of this presentation. How can enzyme catalysis be understood in terms of transition state theory? The simple explanation is that enzyme catalysis lowers the standard Gibbs energy of activation. This can be either an entropic or enthalpic effect. Much more detailed interpretations of enzyme catalysis in terms of transition state theory have been developed. The basic idea is that formation of the enzyme–substrate complex alters the Gibbs energy profile of the reaction such that the Gibbs energy of activation for the reaction is greatly lowered. Essentially, the Gibbs energy change associated with binding of the substrate to the enzyme is used to promote catalysis.

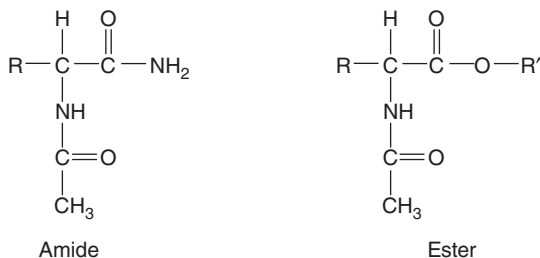
6.3 α -CHYMOTRYPSIN

As an example of how mechanisms can be developed for the action of enzymes, early studies of the enzyme α -chymotrypsin will be discussed. Proteolytic enzymes were among the first to be studied, not because of their intrinsic interest, but because they could easily be isolated in a reasonably pure form. In fact, for many years the availability of large quantities of pure enzyme severely restricted the range of enzymes that could be studied in mechanistic detail. With the ease of cloning and modern expression systems, this is no longer a limitation.

The enzyme α -chymotrypsin has a molecular weight of about 25,000 and consists of three polypeptide chains covalently linked by disulfides. It is an *endopeptidase*; that is, it cleaves peptide bonds in the middle of a protein. Enzymes that cleave peptide bonds at the end of a protein are called *exopeptidases*. The overall reaction can be written as

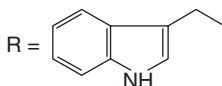


Experimentally, it is found that the enzyme has a strong preference for R being an aromatic group, amino acids phenylalanine, tyrosine, and tryptophan, but hydrophobic amino acids such as isoleucine, leucine, and valine are also good substrates. Studying this reaction with a protein as substrate is very difficult because the molecular structure of the substrate is changing continuously as peptide bonds are cleaved. This makes quantitative interpretation of the kinetics virtually impossible. Consequently, the first step in elucidating the molecular details of the enzyme was to develop “model” substrates, molecules that have the important features of the protein substrate but are much simpler. Chymotrypsin is also a good esterase; that is, it hydrolyzes esters. This proved very useful in elucidating the mechanism of action of the enzyme. The most successful model substrates have the general structure



where R is the aromatic residue associated with phenylalanine, tyrosine, and tryptophan. Note that the amino group of the aromatic amino acid is acetylated. The enzyme will not work on a free amine, as might be expected for an endopeptidase. Furthermore, only L-amino acids are substrates.

We will consider some of the results obtained with tryptophan as the amino acid.



The hydrolysis of tryptophan model substrates follows Michaelis–Menten kinetics, and some of the results obtained for various esters and the amide are shown in Table 6-1. In this table, $k_{\text{cat}} = V_{\text{m}}/[E_0]$, and R' is the group covalently linked to the tryptophan carboxyl. The fact that k_{cat} is essentially the same for all esters suggests that a common intermediate is formed and that the breakdown of the intermediate is the rate determining step in the mechanism. A different mechanism must occur for the amide or there is a different slow step in the mechanism.

An inherent deficiency of steady-state kinetic studies of enzymes is that the enzyme concentration is very low. Consequently, the intermediates in the reaction sequence cannot be detected directly. Limited information about the intermediates can be obtained through steady-state kinetic studies by methods not discussed here, for example, the pH dependence of the kinetic parameters and the use of isotopically labeled substrates that alter the kinetic parameters. However, in order to study the intermediates directly, it is necessary to use sufficiently high enzyme concentrations so that the intermediates can be observed directly. This is the realm of *transient kinetics*. The difficulty for enzymatic reactions is that the reactions become very

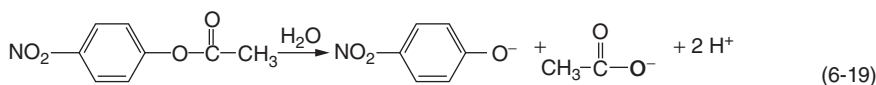
TABLE 6-1. Steady-State Constants for Chymotrypsin.

R'	$k_{\text{cat}} (\text{s}^{-1})$	$K_{\text{M}} (\text{mM})$
Ethyl	27	~5
Methyl	28	~5
<i>p</i> -Nitrophenyl	31	~5
Amide	0.03	~0.09

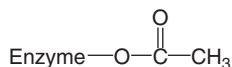
Source: Adapted from R. J. Foster and C. Niemann, *J. Am. Chem. Soc.* **77**, 1886 (1955) and L. W. Cunningham and C. S. Brown, *J. Biol. Chem.* **221**, 287 (1956).

fast, typically occurring in the millisecond and microsecond range. This requires special experimental techniques. Transient kinetic studies have proved invaluable in elucidating how enzymes work.

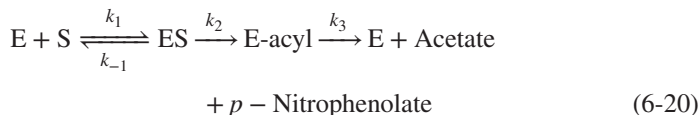
In the case of chymotrypsin, a substrate was sought for which a color change occurred upon reaction in order that the changes in concentration could easily and rapidly be observed. The first such substrate studied was *p*-nitrophenyl acetate, which is hydrolyzed by chymotrypsin:



The advantage of this substrate is that the phenolate ion product is yellow so that the time course of the reaction can easily be followed spectrophotometrically. With this substrate, it was possible to look at the establishment of the steady state, as well as the steady-state reaction. A very important finding was that the reaction becomes extremely slow at low pH. An intermediate, in fact, can be isolated by using a radioactive acetyl group in the substrate. The intermediate is an acylated enzyme. Later studies showed that the acetyl group is attached to a serine residue on chymotrypsin (4).



On the basis of the kinetic studies and isolation of the acyl enzyme, a possible mechanism is the binding of substrate, acylation of enzyme, and deacylation of the enzyme. The elementary steps can be written as



Does this mechanism give the correct steady-state rate law? The rate law can be derived as for the simple Michaelis–Menten mechanism using the steady-state approximation for all of the enzyme species and the mass conservation relationship for enzyme:

$$[\text{E}_0] = [\text{E}] + [\text{ES}] + [\text{E-acyl}] \quad (6-21)$$

$$\begin{aligned} -\frac{d[\text{ES}]}{dt} &= (k_{-1} + k_2)[\text{ES}] - k_1[\text{E}][\text{S}] \approx 0 \\ -\frac{d[\text{E-acyl}]}{dt} &= k_3[\text{E-acyl}] - k_2[\text{ES}] \approx 0 \\ v &= k_3[\text{E-acyl}] \end{aligned} \quad (6-22)$$

When these relationships are combined, it is found that the Michaelis–Menten rate law, Eq. (6-8), is obtained, with

$$V_m = \frac{k_2 k_3 [E_0]}{k_2 + k_3} \quad (6-23)$$

$$K_M = \left(\frac{k_{-1} + k_2}{k_1} \right) \left(\frac{k_3}{k_2 + k_3} \right) \quad (6-24)$$

For ester substrates, the slow step is deacylation of the enzyme, or $k_3 \ll k_2$. In this case, $k_{cat} = k_3$ and $K_M = [(k_{-1} + k_2)/k_1](k_3/k_2)$.

How did the transient kinetic studies contribute to the postulation of this mechanism? It is worth spending some time to analyze the mechanism in Eq. (6-20) in terms of transient kinetics. This analysis can serve as a prototype for understanding how transient kinetics can be used to probe enzyme mechanisms, not only for chymotrypsin but also for other enzyme reactions. The proposed mechanism predicts that the acyl enzyme should accumulate and be directly observable. If we consider only a single turnover of the enzyme at very early times and high substrate concentrations, all of the enzyme should be converted to the acyl enzyme, followed by a very slow conversion of the acyl enzyme to the free enzyme through hydrolysis ($k_3 \ll k_2$). The slow conversion to enzyme will follow Michaelis–Menten kinetics (Eq. (6-13)) initially as long as the total substrate concentration is much greater than the total enzyme concentration. Thus, the complete solution to the rate equations for the mechanism in Eq. (6-20) should contain two terms: (i) the rate of conversion of the enzyme to the acyl enzyme and (ii) the rate of the overall reaction that is limited by the rate of hydrolysis of the acyl enzyme.

To simplify the analysis, let us assume the following: k_3 is approximately zero at early times ($k_2 \gg k_3$); the formation of the initial enzyme–substrate complex, ES, is very rapid relative to the rates of the acylation and deacylation steps so that ES is in a steady state ($d[ES]/dt = 0$); and the total substrate concentration is much greater than the total enzyme concentration. The last assumption is similar to that for the steady-state approximation used earlier for the Michaelis–Menten mechanism. However, in this case, the enzyme concentration is sufficiently high so that the concentration of the intermediate can be detected. The rate of formation of the intermediate is given by

$$\frac{d[E\text{-acyl}]}{dt} = k_2[ES] \quad (6-25)$$

Because the phenolate ion is formed when the acyl enzyme is formed, the rate of formation of phenolate ion is the same as the rate of formation of the acyl enzyme. This is actually a measurement of the rate of establishment of the steady state. In order to integrate this equation, we make use of the identity

$$[E] + [ES] = [E] \left(1 + \frac{[E]}{[ES]} \right) = [ES] \left\{ 1 + \frac{(k_{-1} + k_2)}{k_1[S]} \right\}$$

and mass conservation

$$[E_0] = [E] + [ES] + [E\text{-acyl}]$$

or

$$[E] + [ES] = [E_0] - [E\text{-acyl}]$$

Insertion of these relationships into Eq. (6-25) gives

$$\frac{d[E\text{-acyl}]}{dt} = \frac{k_2([E_0] - [E\text{-acyl}])}{1 + \frac{(k_{-1} + k_2)}{k_1[S]}} \quad (6-26)$$

Integration of this equation gives

$$[E\text{-acyl}] = [E_0](1 - e^{-k'_2 t}) \quad (6-27)$$

with $k'_2 = k_2 / \{1 + (k_{-1} + k_2) / (k_1[S])\}$.

Note that when $t = 0$, $[E\text{-acyl}] = 0$, and when $t = \infty$, $[E\text{-acyl}] = [E_0]$, as expected. The time dependence of phenolate, P_1 , formation at early times is

$$[P_1] = v t + [E_0](1 - e^{-k'_2 t}) \quad (6-28)$$

The time course of the reaction as described by Eq. (6-28) is shown schematically in Figure 6-3. The second term in the equation dominates at early times, and the exponential rise can be analyzed to give k'_2 . The linear portion of the curve corresponds to the Michaelis–Menten initial velocity and can be used to obtain V_m and K_M . The maximum velocity for this limiting case is $k_3[E_0]$, so that both k_2 and k_3 can be determined. A more exact analysis that does not assume $k_3 = 0$ is a bit more complex, but the end result is similar. The rate equation contains the same two terms as in Eq. (6-28). However, the second term is now more complex: the exponent is $(k'_2 + k_3)t$ instead of $k'_2 t$; and the amplitude is $[E_0][k'_2 / (k'_2 + k_3)]^2$ rather than $[E]_0$. The first term is the same, $v t$, with the maximum velocity and Michaelis constant defined by Eqs. (6-23) and (6-24).

Finally, we consider the situation when the rate of hydrolysis of the acyl enzyme is very fast relative to its formation ($k_3 \gg k_2$). We have seen previously that fast steps occurring after the rate-determining step do not enter into the rate law. Therefore, the mechanism is equivalent to the Michaelis–Menten mechanism with only a single intermediate (Eq. (6-2)), with the rate-determining step being the acylation of the enzyme ($V_m = k_2[E_0]$). If the transient kinetics are analyzed, the rate is still given by Eq. (6-28), but the amplitude of the second term is essentially zero, rather than $[E_0]$. Thus, the transient kinetics and steady-state kinetics provide the same information. The presence of an initial “burst” of product, as shown in Figure 6-3, is commonly used as a diagnostic for the presence of an intermediate. If the burst is not present, that is, the straight line portion of the curve extrapolates through zero, this means either that an intermediate is not formed, or that its rate of disappearance is much

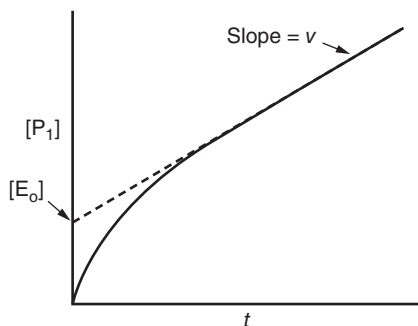


FIGURE 6-3. Schematic representation for the kinetics of an enzymatic reaction displaying “burst” kinetics. The product of the reaction, P_1 , is plotted versus the time. As described by Eq. (6-28) for the limiting case of $k_2 \gg k_3$, the slope of the linear portion of the curve is the steady-state initial velocity, v , and the initial exponential time dependence is characterized by the rate constant k'_2 and a “burst” amplitude of $[E_0]$.

faster than its rate of appearance. This exhaustive, perhaps exhausting, discussion of the mechanism in Eq. (6-20) is a good example of what must be done in order to arrive at a mechanism for an enzyme reaction by kinetic studies. Both steady-state kinetics and transient kinetics are necessary for a complete picture.

The proposed mechanism accounts for all of the facts presented: the observed rate law, the same k_{cat} for all ester substrates, and a common intermediate. What about amides? For amides, it turns out that the slow step is acylation of the enzyme, that is, $k_2 \ll k_3$, so that $k_{\text{cat}} = k_2$ and $K_M = (k_{-1} + k_2)/k_1$. As discussed earlier, if the acylation of the enzyme is slow relative to deacylation, $k_3 \gg k_2$, the intermediate does not accumulate and therefore cannot be observed directly. Literally, hundreds of experiments are consistent with this mechanism—a very remarkable achievement. Additional experiments that permitted the very fast reactions prior to the acylation of the enzyme to be studied show that this mechanism is too simple. At least one additional elementary step is required. The sequential elementary steps are (i) binding of substrate to the enzyme, (ii) a conformational change of the enzyme–substrate complex, (iv) acylation of the enzyme, and (v) deacylation of the enzyme.

The ultimate goal of kinetic studies is to understand the mechanism in terms of molecular structure. In the case of chymotrypsin, the three-dimensional structure of the enzyme is known and is shown in Figure 6-4a. A well-defined binding pocket is observed for the aromatic side chain of the specific substrates of the enzyme. The pocket is lined with nonpolar side chains of amino acids and is very hydrophobic. Up to three amino acids coupled to the N-terminus of the aromatic amino acid interact with a short range of antiparallel β -sheet in the enzyme. A hydrophobic site is also observed for the amino acid attached to the C-terminus of the substrate. This explains why a free carboxyl group cannot bind and therefore why chymotrypsin is an endopeptidase. As shown in Figure 6-4b, a triad of amino acids, serine-195, histidine-57, and aspartate-102, is observed in the active site region and is found in other serine proteases. The imidazole acts as a general base during the nucleophilic attack of the serine hydroxyl on the substrate. A tetrahedral intermediate probably

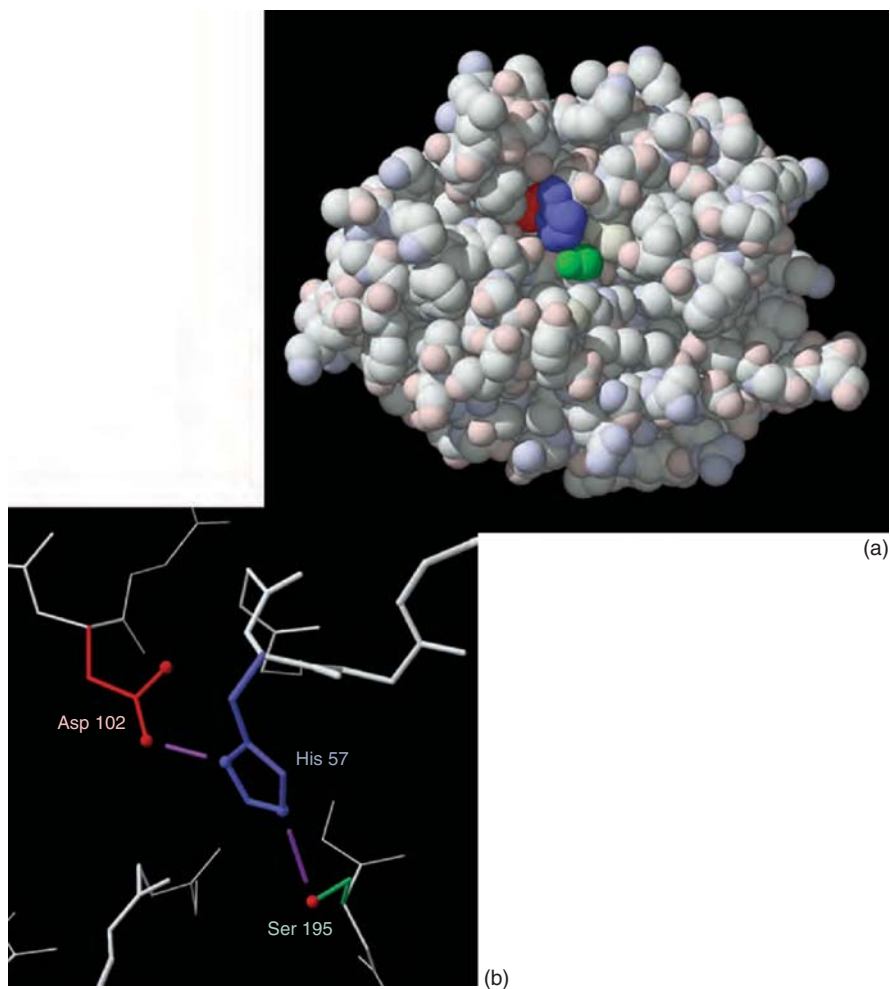


FIGURE 6-4. (a) Space filling representation of chymotrypsin, showing the active site pocket where substrate binding occurs. The blue at the active site is imidazole, the green is serine, and the red is aspartate. (b) Arrangement of the protein residues at the active site of chymotrypsin. The serine is the residue acylated; the imidazole of the histidine serves as a general base catalyst; and the aspartate carboxyl group is hydrogen bonded to the imidazole. Copyright by Professor Jane Richardson. Reprinted with permission.

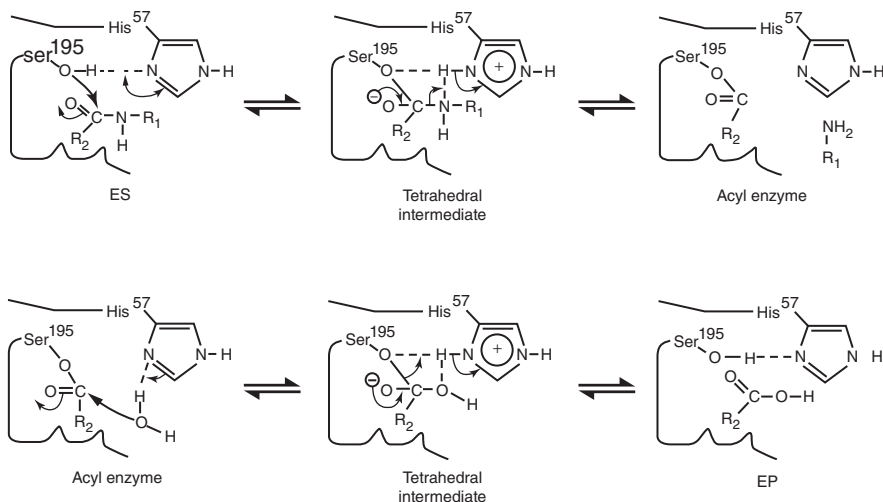


FIGURE 6-5. A mechanism for the hydrolysis of peptides or amides by chymotrypsin. The imidazole acts as a general base to assist the nucleophilic attack of serine on the substrate or the nucleophilic attack of water on the acyl enzyme.

is formed prior to acyl enzyme formation. Water serves as the nucleophile for the hydrolysis of the acyl enzyme, with imidazole again participating as the general base. An abbreviated version of the catalytic mechanism is shown in Figure 6-5.

This greatly truncated story of the elucidation of the mechanism of action of chymotrypsin illustrates several important concepts: the steady-state approximation, the use of steady-state kinetics in determining chemical mechanisms, the importance of transient kinetics in elucidating intermediates in a mechanism, and the interpretation of the mechanism in terms of molecular structure. More complete discussions of chymotrypsin are available (1,2,5).

6.4 PROTEIN TYROSINE PHOSPHATASE

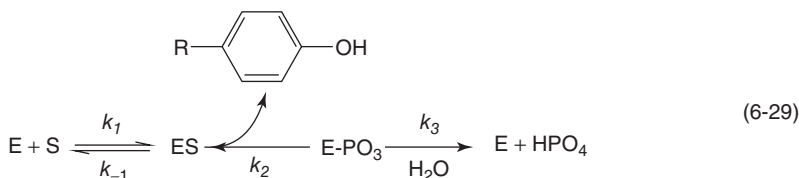
In our earlier discussion of metabolism, we saw that the energy obtained in glycolysis is stored as a phosphate ester in ATP and serves as a source of free energy for biosynthesis. The importance of phosphate esters is not confined to this function. Phosphorylation and dephosphorylation of proteins play a key role in signal transduction and the regulation of many cell functions. For example, the binding of hormones to cell surfaces can trigger a cascade of such reactions that regulate metabolism within the cell. The cell processes modulated by this mechanism include T-cell activation, the cell cycle, DNA replication, transcription and translation, and programmed cell death, among many others. Both the kinases, responsible for protein phosphorylation, and the phosphatases, responsible for protein dephosphorylation, have been studied

extensively. We shall restrict this discussion to the phosphatases, and primarily to one particular type of phosphatase.

The phosphorylation of proteins primarily occurs at three sites, namely, the side chains of threonine, serine, and tyrosine. Protein phosphatases can be divided into two structural classes: the serine/threonine-specific phosphatases that require metal ions and tyrosine phosphatases that do not require metals but use a nucleophilic cysteine to cleave the phosphate. The latter class includes dual function phosphatases that can hydrolyze phosphate esters on serine and threonine, as well as tyrosine. Many reviews of these enzymes are available (cf. Refs. 6 and 7). We will discuss the mechanism of a particular example of tyrosine-specific phosphatase, namely, an enzyme found in both human and rat that has been extensively studied structurally and mechanistically (6).

The overall three-dimensional structure of the catalytic domain of the enzyme is shown in Figure 6-6a, and the active site region is shown in Figure 6-6b with vanadate bound to the catalytic site. Vanadate is frequently used as a model for phosphate esters because the oxygen bonding to vanadium is similar to the oxygen bonding to phosphorus, and vanadate binds tightly to the enzyme. A cysteine rests at the base of the active-site cleft. An aspartic acid and a threonine (or serine in similar enzymes) are also near the active site and are postulated to play a role in catalysis, as discussed in the subsequent sections.

The kinetic mechanism that has been postulated is the binding of the protein substrate followed by transfer of the phosphoryl group from tyrosine to the cysteine at the active site, and finally hydrolysis of the phosphorylated cysteine:



As with chymotrypsin, a protein is not a convenient substrate to use because of the difficulty in phosphorylating a single specific tyrosine and the lack of a convenient method for following the reaction progress. In this case, the model substrate used was *p*-nitrophenyl phosphate. Similar to chymotrypsin, when the enzyme is phosphorylated the yellow phenolate ion is released so that the reaction progress can easily be monitored. Extensive steady-state and transient kinetic studies have been carried out, and we will only present a few selected results that summarize the findings (8,9).

The role of the cysteine was firmly established by site-specific mutagenesis, as converting the cysteine to a serine results in an inactive enzyme. The cysteine is postulated to function as a nucleophile with the formation of a cysteine phosphate. However, when transient kinetic studies were carried out with the native enzyme, a burst phase was not observed. This is shown in Figure 6-7a. As we noted previously, this indicates that either a phosphoenzyme intermediate is not formed or its hydrolysis is much faster than its formation. The aspartic acid and serine or threonine shown at the active site of the enzyme in Figure 6-6 are conserved in many

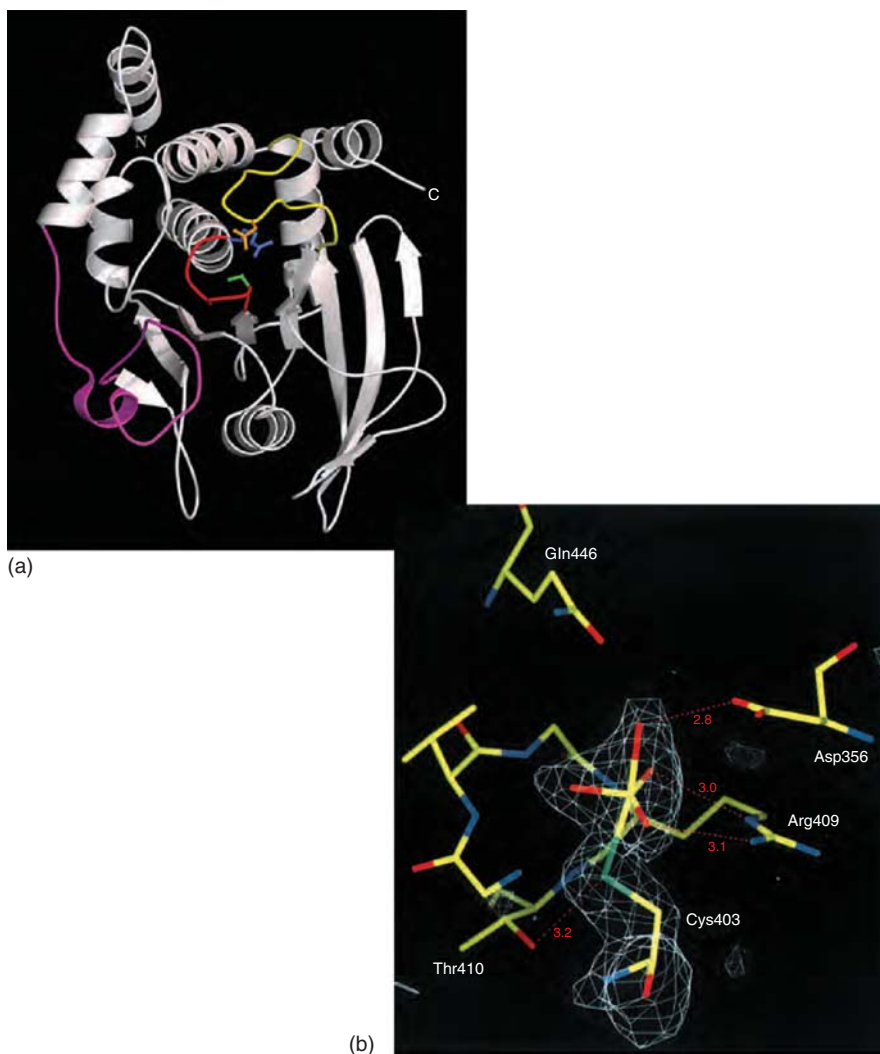


FIGURE 6-6. (a) Backbone representation of the catalytic domain of protein tyrosine phosphatase. Coils and arrows represent α -helices and β -strands, respectively. The cysteine (green), arginine (red), and aspartate (gold) are also shown. Reproduced from J. M. Denu, J. A. Stuckey, M. A. Saper, and J. E. Dixon, *Cell* **87**, 316 (1996). Copyright © 1996 with permission from Elsevier. (b) Vanadate bound at the active-site cysteine of protein tyrosine phosphatase. Other amino acids at the catalytic site are indicated, including the threonine and aspartate that participate in the catalytic mechanism. Reproduced from J. M. Denu, D. L. Lohse, J. Vijayalakshmi, M. A. Saper, and J. E. Dixon, *Proc. Natl. Acad. Sci. USA* **93**, 2493 (1996). Reprinted with permission of the Proceedings of the National Academy of Sciences USA. Reproduced by permission of the publisher via Copyright Clearance Center, Inc.

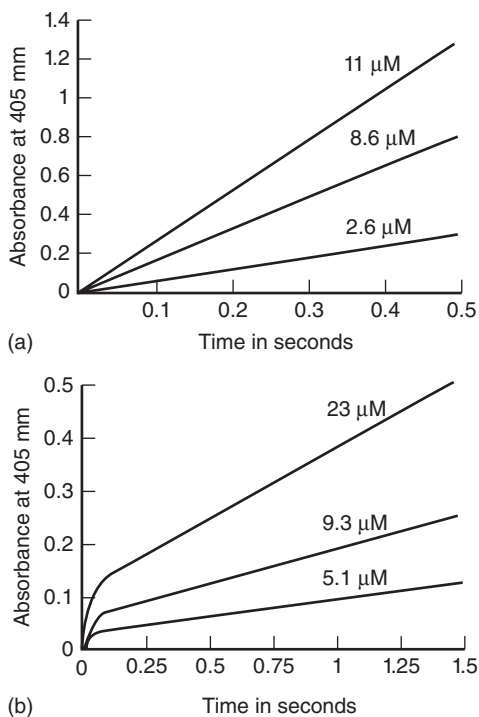


FIGURE 6-7. Transient kinetic time course for the hydrolysis of *p*-nitrophenyl phosphate by native (a) and mutant (b) protein tyrosine phosphatase. The point mutation substituted alanine for serine 222. The concentrations of substrate are shown next to the traces. Reprinted with permission from D. L. Lohse, J. M. Denu, N. Santoro, and J. E. Dixon, *Biochemistry* **36**, 4568 (1997). Copyright © 1997 American Chemical Society.

different enzymes of this class of tyrosine phosphatases. Consequently, it was decided to mutate these residues. If the serine is mutated to alanine, the enzyme is still active, but now a burst is observed, as shown in Figure 6-7b. This burst was postulated to represent the formation of the phosphoenzyme, and the rate constants were obtained for the phosphorylation of cysteine and its hydrolysis. The results of the transient kinetic experiments are presented in Table 6-2. For the native enzyme, only k_{cat} can be determined and is equal to k_2 . For the serine to alanine mutant, the rate constant for hydrolysis of the intermediate is the same as k_{cat} , as expected. This mutation also suggests that serine plays a role in the hydrolysis mechanism.

The importance of the aspartic acid in the mechanism was inferred not only from its conservation in many different enzymes, but also by the pH dependence of the steady-state kinetic parameters, which suggested it functioned as a general acid in the phosphorylation of the enzyme. When the aspartic acid was mutated to asparagine, the enzyme still functioned, but k_{cat} was only about 1% of that of the native enzyme (Table 6-2). The transient kinetics did not show a burst phase, indicating that formation of the phosphoenzyme was still rate determining. As expected, the pH dependence of the steady-state kinetic parameters was altered by this mutation.

TABLE 6-2. Kinetic Parameters for Native and Mutated Tyrosine Phosphatases

Enzyme	k_{cat} (s^{-1})	k_2 (s^{-1})	k_3 (s^{-1})	$K(\mu\text{M})^a$
Native	20			
Serine 222 $\rightarrow$ alanine	1.3	34	1.4	0.46
Aspartate 181 $\rightarrow$ asparagine	0.27			
Serine 222 $\rightarrow$ alanine and Aspartate 181 $\rightarrow$ asparagine	0.055	1.9	0.057	0.81

Source: Adapted from D. L. Lohse, J. M. Denu, N. Santoro, and J. E. Dixon, *Biochemistry* **36**, 4568 (1997).

^a $(k_{-1} + k_2)/k_1$.

Finally, the double mutated enzyme was prepared in which the serine was changed to alanine and the aspartate to asparagine. The transient kinetics showed a burst so that the slow step was now the hydrolysis of the intermediate, as with the single mutant in which serine was changed to alanine. However, k_{cat} was reduced even further, and the rate constants for both phosphorylation of the enzyme and hydrolysis were greatly reduced (Table 6-2). On the basis of these and other data, the aspartate was postulated to serve as a general base in the hydrolysis of the intermediate. The intermediate is sufficiently stable in the double mutant so that it could be observed directly with phosphorus nuclear magnetic resonance (8). The mechanism proposed for this reaction has two distinct transition states, one for phosphorylation of the cysteine and another for hydrolysis of the phosphorylated enzyme. Based on the experiments discussed and other data, especially structural data, structures have been proposed for the two transition states as shown in Figure 6-8.

The mechanism of tyrosine phosphatases illustrates several important points. First, the general usefulness of the analysis developed for chymotrypsin is apparent. Both transient and steady-state kinetic experiments were important in postulating a mechanism. Second, the importance of site-specific mutations in helping to establish the mechanism is evident. However, this method is not without pitfalls. It is important to establish that mutations do not alter activity through structural changes in the molecule. In the present case, experiments were carried out to establish the structural integrity of the mutant enzymes. This ensures that the mutations are altering only the chemical aspects of the mechanism. Finally, this discussion again illustrates the many different types of experiments that must be performed in developing mechanisms. The kinetic results must be bolstered by structural and chemical information. Many other studies of enzymes could profitably be discussed, but instead we will turn our attention to kinetic studies in other types of biological systems.

6.5 RIBOZYMES

Enzymes are the most efficient and prevalent catalysts in physiological systems, but they are not the only catalysts of biological importance. RNA molecules have also been found to catalyze a wide range of reactions. Most of these reactions involve the processing of RNA, cutting RNA to the appropriate size or splicing RNA. RNA

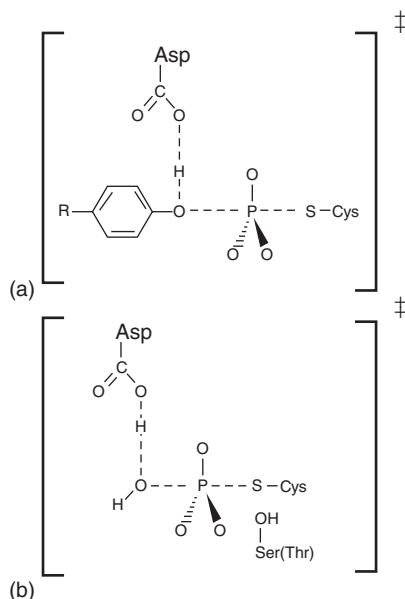


FIGURE 6-8. Proposed transition states for the protein tyrosine phosphatase reaction. (a) Formation of the cysteine phosphate intermediate with a tyrosine phosphate as substrate. (b) Hydrolysis of the cysteine phosphate intermediate. Adapted from J. M. Denu, D. L. Lohse, J. Vijayalakshmi, M. A. Saper, and J. E. Dixon, *Proc. Natl. Acad. Sci. USA* **93**, 2493 (1996). Republished with permission of the National Academy of Sciences USA.

has also been implicated in peptide bond formation on the ribosome and has been shown to hydrolyze amino acid esters. These catalytic RNAs are called *ribozymes*. They are much less efficient than a typical protein enzyme and sometimes catalyze only a single event. In the latter case, this is not true catalysis. Ribozymes also may work in close collaboration with a protein in the catalytic event. We will consider an abbreviated discussion of a ribozyme, ribonuclease P. For more information, many reviews of ribozymes are available (10–12).

Ribonuclease P catalyzes an essential step in tRNA maturation, namely, the cleavage of the 5' end of a precursor tRNA (pre-tRNA) to give an RNA fragment and the mature tRNA. This reaction is shown schematically in Figure 6-9. Also shown in this figure is a representation of the catalytic RNA. The catalytic RNA is about 400 nucleotides long and catalyzes the maturation reaction by itself *in vitro*. *In vivo*, a protein of about 120 amino acids also participates in the catalysis. The precise role of the protein is not known, but it appears to alter the conformation of the RNA. The protein–RNA enzyme has a broader selectivity for biological substrates and is a more efficient catalyst (13). Metal ions are also involved in this reaction, with Mg^{2+} being the most important physiologically. Metal ions probably play a role both in catalysis and in maintaining the active conformation of the RNA, but we will not consider this aspect of the reaction here. Transient kinetic studies of the ribozyme have been carried out, and the minimal mechanism consists of the binding of pre-tRNA to the

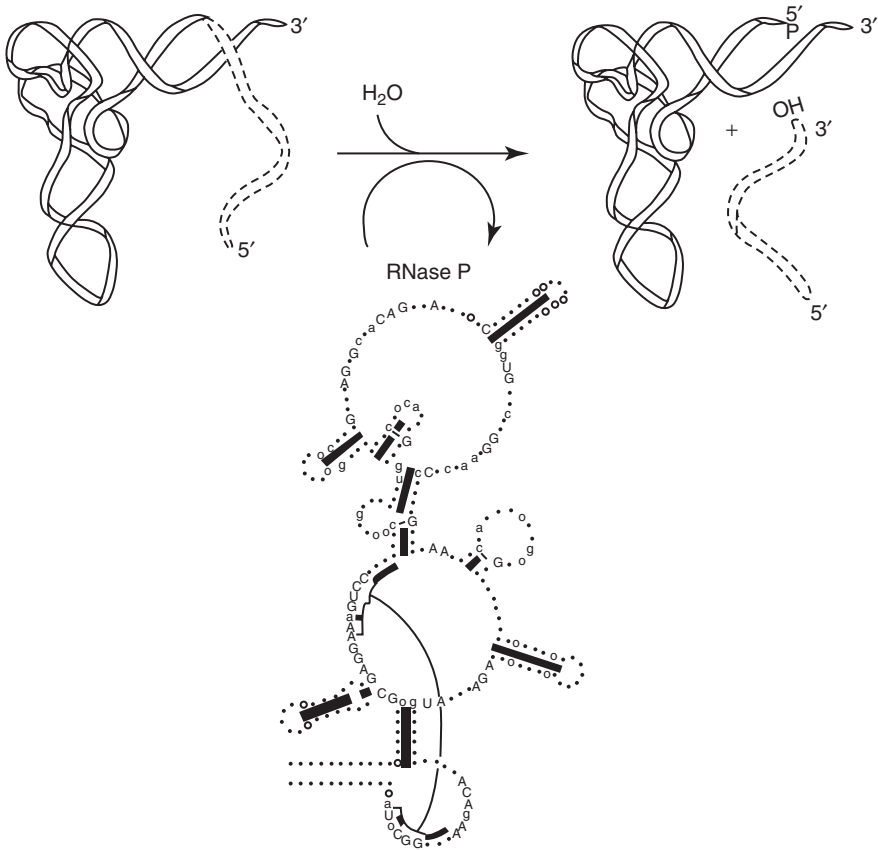


FIGURE 6-9. The reaction of pre-tRNA catalyzed by ribonuclease P is shown schematically at the top of the figure. The 5' leader sequence of the pre-tRNA is removed in the reaction. The eubacterial consensus structure of ribonuclease P RNA is shown in the lower part of the figure. Proven helices are designated by filled rectangles, invariant nucleotides by uppercase letters, >90% conserved nucleotides by lowercase letters, and less conserved nucleotides by dots. Reproduced from J. W. Brown and N. R. Pace, *Nucleic Acids Res.* **20**, 1451 (1992). Reproduced by permission of Oxford University Press.

ribozyme, cleavage of the phosphodiester bond, and independent dissociation of both products. We will present some of the results obtained with the RNA component of ribonuclease P.

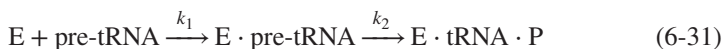
If transient experiments are carried out with ribonuclease P with excess substrate, a “burst” of the tRNA product occurs at short times, followed by an increase in the product concentration that is linear with time. This behavior is familiar by now, as it has been observed for chymotrypsin and protein tyrosine phosphatase, as discussed earlier. In the case of ribonuclease P, a covalent intermediate is not formed with the ribozyme; instead, the dissociation of products is very slow relative to the hydrolytic

reaction. The mechanism can be written as



where E is RNAase P RNA and P is the pre-tRNA fragment product. This is not exactly the same as the mechanism discussed for chymotrypsin, but the analysis is similar if the second step is assumed to be much slower than the first. (The detailed analysis is given in Ref. 14.) There is an important lesson to be learned here. “Burst” kinetics are observed whenever the product is formed in a rapid first step followed by regeneration of the enzyme in a slower second step. The apparent rate constants for these two steps can be determined.

What is desired, however, is to distinguish between the binding of enzyme and substrate and the hydrolysis; both of these reactions are aggregated into the first step in Eq. (6-30). In the case of chymotrypsin and protein tyrosine phosphatase, we assumed that the binding step was rapid relative to the formation of the enzyme–substrate intermediate. In this case, we cannot assume that binding is rapid relative to hydrolysis. The two steps were resolved by single turnover experiments: an excess of enzyme was used and because dissociation of products is slow, only a single turnover of substrate was observed (14). The mechanism can be written as



The results obtained are shown in Figure 6-10 for two different concentrations of enzyme. At the lower concentration of enzyme, it is clear that the curve is sigmoidal, rather than hyperbolic. This is because the first step is sufficiently slow that it takes some time for the concentration of the first enzyme–substrate complex to build up.

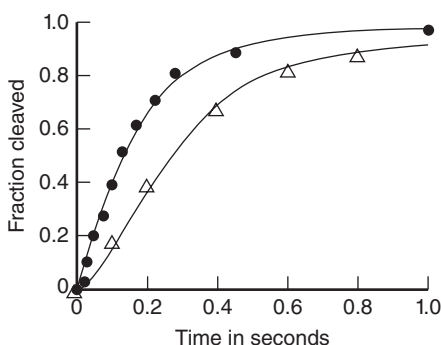


FIGURE 6-10. Single turnover measurements of the hydrolysis of pre-tRNA catalyzed by the RNA component of RNAase P. The fraction of pre-tRNA cleaved is plotted versus time. The pre-tRNA was mixed with excess concentrations of RNAase P RNA, 1.4 μM (Δ), or 19 μM ($\bullet$). The data are the fit to a mechanism of two consecutive first-order reactions (Eq. (6-38)). Reprinted with permission from J.A. Beebe and C.A. Fierke, *Biochemistry* **33**, 10294 (1994). Copyright 1994 American Chemical Society.

The data at the higher enzyme concentration are essentially hyperbolic because the rate of the first reaction is greater at higher enzyme concentration. These data can be explained by analyzing the mechanism in Eq. (6-31) as two consecutive irreversible first-order reactions. Why irreversible? Fortunately, the rate of dissociation of pre-tRNA from the ribozyme is sufficiently slow that it does not occur on the time scale of the experiment, and furthermore, the analysis of the product gives both E-tRNA and free tRNA so that if some dissociation of tRNA occurs it is not relevant. Why is the first step assumed to be first order? This is another example of a pseudo-first-order rate constant because the enzyme concentration is effectively constant throughout the experiment. The pseudo-first-order rate constant is $k_1[E]$, where k_1 is the second-order rate constant for the reaction of the enzyme with the substrate.

To simplify the nomenclature, we will rewrite the mechanism as



The rate equation for the time dependence of A is

$$-\frac{d[A]}{dt} = k_1[A] \quad (6-33)$$

which is easily integrated to give

$$[A] = [A]_0 e^{-k_1 t} \quad (6-34)$$

where $[A]_0$ is the starting concentration, in this case of pre-tRNA. The rate equation for the time dependence of B is

$$\begin{aligned} -\frac{d[B]}{dt} &= k_2[B] - k_1[A] \\ -\frac{d[B]}{dt} &= k_2[B] - k_1[A]_0 e^{-k_1 t} \end{aligned} \quad (6-35)$$

The solution to this differential equation is

$$[B] = \frac{[A]_0 k_1}{k_2 - k_1} (e^{-k_1 t} - e^{-k_2 t}) \quad (6-36)$$

This solution does not work if $k_1 = k_2$ as the denominator goes to zero. For this special case,

$$[B] = k_1 [A]_0 t e^{-k_1 t} \quad (6-37)$$

Finally, the time dependence of C can be obtained from mass balance since $[A]_0 = [A] + [B] + [C]$:

$$[C] = [A]_0 \left(1 - \frac{k_2}{k_2 - k_1} e^{-k_1 t} + \frac{k_1}{k_2 - k_1} e^{-k_2 t} \right) \quad (6-38)$$

If Eq. (6-38) is used to analyze the data in Figure 6-10, it is found that $k_1 = 6 \times 10^6 \text{ M}^{-1} \text{ s}^{-1}$ $[\text{RNA}]_0$ and $k_2 = 6 \text{ s}^{-1}$. Understanding how to analyze “burst” kinetics and consecutive first-order reactions is sufficient for the kinetic analysis of many enzymatic reactions, as the conditions can usually be adjusted to conform to these relatively simple mechanisms.

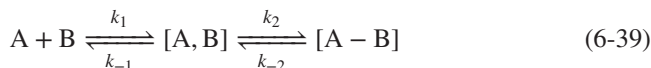
The mechanistic work carried out with RNaseP has permitted the establishment of a minimal mechanism for the RNA portion of the enzyme. However, much remains to be done. The roles of metals, specific groups on the RNA, and the protein remain to be delineated. Understanding how ribozymes function is at the forefront of modern biochemistry and has important implications for both physiology and the evolution of enzymes.

6.6 DNA MELTING AND RENATURATION

We will conclude this chapter with a discussion of the denaturation and renaturation of DNA. Understanding the dynamics of such processes is clearly of biological importance. At the outset, it must be stated that a detailed understanding of the kinetics and mechanisms has not been achieved. However, this is not for lack of effort, and a qualitative understanding of the mechanisms has been obtained.

We will begin with some of the elementary steps in the dynamics of the interactions between the two chains of helical DNA. As discussed in Chapter 3, the thermodynamics of hydrogen bonding between bases has been studied in nonaqueous solvents, where the dimers formed are reasonably stable. Kinetic studies of hydrogen-bonded dimers have also been carried out, and the reactions have been found to be extremely fast, occurring on the submicrosecond and nanosecond time scales. For example, the kinetics of formation of a hydrogen-bonded dimer between 1-cyclohexyluracil and 9-ethyladenine has been studied in chloroform (15). (The hydrocarbon arms have been added to the bases to increase their solubility in chloroform.) The reaction is found to occur in a single step with a second-order rate constant for the formation of the dimer of $2.8 \times 10^9 \text{ M}^{-1} \text{ s}^{-1}$ and a dissociation rate constant of $2.2 \times 10^7 \text{ s}^{-1}$ at 20°C . The second-order rate constant is the maximum possible value; that is, it is the value expected if every collision between the reactants produced a hydrogen-bonded dimer. The upper limit for the rate constant of a bimolecular reaction can be calculated from the known rates of diffusion of the reactants in the solvent. In all cases where hydrogen-bonded dimer formation has been studied, the formation of the dimer has been found to be diffusion controlled.

What does this tell us about the rate of hydrogen bond formation that occurs after the two reactants have diffused together? To answer this question, we will postulate a very simple mechanism, namely, diffusion together of the reactants to form a dimer that is not hydrogen bonded, followed by the formation of hydrogen bonds. This can be written as



In this mechanism, k_1 is the rate constant for the diffusion-controlled formation of the intermediate and k_{-1} is the rate constant for the diffusion-controlled dissociation of the intermediate. Since only a single step is observed in the experiments, assume that the initial complex formed is in a steady state:

$$[A, B] = \frac{k_1[A][B]}{(k_{-1} + k_2)} + \frac{k_{-2}[A - B]}{(k_{-1} + k_2)} \quad (6-40)$$

The overall rate of the reaction is

$$\frac{d[A - B]}{dt} = k_2[A, B] - k_{-2}[A - B] \quad (6-41)$$

If Eq. (6-40) is substituted into Eq. (6-41), we obtain

$$\frac{d[A - B]}{dt} = k_f[A][B] - k_r[A - B] \quad (6-42)$$

with

$$k_f = \frac{k_1 k_2}{(k_{-1} + k_2)}$$

and

$$k_r = \frac{k_{-1} k_{-2}}{(k_{-1} + k_2)}$$

The experimental results indicate that $k_f = k_1$. This is true if $k_2 \gg k_{-1}$ or formation of the hydrogen bonds is much faster than diffusion apart from the intermediate. However, we can calculate the value of k_{-1} ; it is about 10^{10} s^{-1} . Therefore, the rate constant for formation of the hydrogen bonds, k_2 , must be greater than about 10^{11} s^{-1} . A more exact analysis would put in a separate step for formation of each of the two hydrogen bonds, but this would not change the conclusion with regard to the rate of hydrogen bond formation. Hydrogen bond formation is very fast!

We know that hydrogen bonding alone cannot account for the stability of the DNA double helix: Base stacking and hydrophobic interactions also are important. An estimate of the rate of base stacking formation and dissociation has been obtained in studies of poly A and polydA (16). These molecules undergo a transition from “stacked” to “unstacked” when the temperature is raised. This transition is accompanied by spectral changes and can easily be monitored. Conditions were adjusted so that the molecules were in the middle of the transition, and the temperature was raised very rapidly ($<1 \mu\text{s}$) with a laser. This is an example of a kinetic study near equilibrium. As the system returns to equilibrium at the higher temperature, the relaxation time can be measured and is the sum of the rate constants for stacking and unstacking (Eq. 5-43). The results obtained indicate that the rate constants are in the range of 10^6 – 10^7 s^{-1} . Although this is a very fast process, it is considerably slower than hydrogen bonding or simple rotation of the bases.

Moving up the ladder of complexity, we now consider the formation of hydrogen-bonded dimers between oligonucleotides (17,18). Only a few prototype

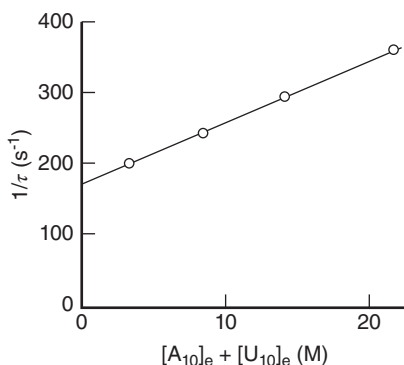


FIGURE 6-11. A plot of $1/\tau$ versus the sum of the equilibrium concentrations of A_{10} and U_{10} at 17°C (Eq. (6-43)). The two reactants combine to form a double-helix structure. Adapted from D. Pörschke and M. Eigen, Cooperative non-enzymic base recognition. III, *J. Mol. Biol.* **62**, 361 (1971). Copyright 1971, with permission from Elsevier.

reactions will be considered, namely, the reaction of a series of A_n oligonucleotides with U_n oligonucleotides to form double-helical dimers. These reactions occur in the millisecond to second time regime with $n = 8\text{--}18$. In all cases, only a single relaxation time was observed. The relaxation time was consistent with a simple bimolecular reaction (Eq. 5-47):

$$\frac{1}{\tau} = k_1([A_n]_e + [U_n]_e) + k_{-1} \quad (6-43)$$

A typical plot of $1/\tau$ versus the sum of the equilibrium concentration is given in Figure 6-11, and some of the rate constants obtained are given in Table 6-3. The bimolecular rate constants are all about $10^6 \text{ M}^{-1} \text{ s}^{-1}$, whereas the dissociation rate constants vary widely, reflecting the stability of the double helix that is formed. Although the second-order rate constant is quite large, it is considerably below the value expected for a diffusion controlled reaction.

A detailed analysis of the mechanism can be made as above for the case of the formation of a hydrogen-bonded dimer. This case is considerably more complex, as a

TABLE 6-3. Rate Constants for the Reaction of $[A_n]$ with $[U_n]$ at 17°C

n	$10^6 k_1 (\text{M}^{-1} \text{s}^{-1})$	$k_{-1} (\text{s}^{-1})$
8	1.7	2400
9	7.1	640
10	0.83	175
11	0.58	28
14	1.44	1

Source: Adapted from D. Pörschke and M. Eigen, *J. Mol. Biol.* **62**, 361 (1971).

separate step is needed for the pairing of each of the n bases. The analysis carried out was bolstered by determination of the activation energies for the rate constants. The conclusion reached is that *nucleation* of the double helix after the molecules have diffused together is rate determining. The formation of the first base pair is rapid, but dissociation is even more rapid. This is also true for the formation of two base pairs. However, when the third base pair is formed, the structure is stabilized. After a nucleus of three base pairs is present, each later base pair forms with a specific rate constant of about 10^7 s^{-1} . Note that this is consistent with the results obtained for the “stacking” and “unstacking” of poly A. Altogether these results indicate that the elementary steps in the dynamics of DNA base pairing and stacking (and whatever other types of interactions are important for base pairing) are very rapid. We now turn to consideration of DNA itself.

As might be expected, the melting and reformation of DNA is very complex. DNA is very heterogeneous, and we have discussed in Chapter 3 that A–T-rich regions are much less thermodynamically stable than G–C-rich regions. In addition, DNA has a complex structure, with loops and bulges. Native DNA has a very precise alignment of base pairs. If extensive melting occurs, base pairings other than the native pairings may form when the temperature is lowered to form the native structure, thereby slowing down reformation of the native structure. The ultimate form of melting is separation of the individual strands. Because the strands are very long, the rate of melting could be limited by the hydrodynamics of moving the chains apart. Bearing these factors in mind, let us examine some experimental results. Fortunately, the melting and reformation can easily be followed because the spectrum of DNA changes when the bases are stacked. (This effect is discussed further in Chapter 10.) Furthermore, the melting/reformation processes are easy to initiate by changing the temperature.

Some typical results for the denaturation of a viral DNA are shown in Figure 6-12. The denaturation was initiated by the raising of the temperature to give a loss in the native spectrum of 25%, 53%, and 95%. In the first case, the loss in absorbance is approximately first order with a rate constant of about 1 min^{-1} . The other two cases do not strictly follow first-order kinetics; that is, the logarithmic plot in Figure 6-12 is not linear, and the rates are clearly much slower. The approximate first-order rate constants at long times are shown in the figure. In contrast to the results in model systems, the melting kinetics cannot be fit to a single relaxation time, and the time scale is 100–1000 s. The viral DNA studied has about 2×10^5 base pairs. If only the elementary steps discussed earlier were involved, the DNA would be completely melted in less than 1 second. How can these results be explained? The rate of melting is governed by the rate of unwinding of the double helix, and the rate of unwinding varies as the melting proceeds. Melting will usually start in the interior of the DNA, and unwinding can only occur if one end of the loop rotates with respect to the other. Initially, this rotation is fast because the double helix is approximately a rod, but as melting proceeds, the bulky loops produced make it harder for one end of the helix to rotate with respect to the other. The rate then becomes slower, as observed. Thus, strand unwinding and separation become rate limiting. In support of this mechanism, if DNA is subjected to a very drastic denaturing force such as a very high pH, all of the base pairs rapidly break in times less than 0.01 s. This does not involve unwinding

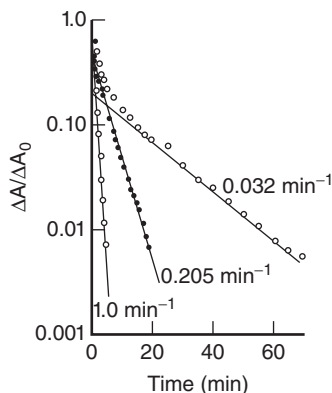


FIGURE 6-12. The kinetics of T2 DNA denaturation in 0.08 M NaCl, 80% formamide. The fraction of the absorbance at 259 nm associated with the native structure versus the time is plotted. The reaction was initiated by heating the three samples to a loss in the native ultraviolet spectrum of 25% (left line), 53% (middle line), and 95% (right line). Reproduced from M. T. Record and B. H. Zimm, *Biopolymers* 11, 1435 (1972). Copyright © 1972 *Biopolymers*. Reprinted by permission of John Wiley & Sons, Inc.

and strand separation as the native DNA can rapidly be formed by lowering the pH. A detailed description of this mechanism has been developed (19).

The slow unwinding time could be a severe problem for the replication of large DNAs. However, nature has solved this problem by creating enzymes that (i) can make breaks in the DNA chains, thereby reducing the length of DNA that must be rotated during the unwinding, and (ii) can unwind DNA.

As a final example, we consider the renaturation of separated DNA strands. DNA can be denatured by raising the temperature significantly above the midpoint of the melting transition or by raising the pH. After the strands have separated, native DNA can be formed by lowering the temperature or lowering the pH. As might be expected, the renaturation process is second order. The rate of reaction depends on the source of the DNA and its length. The dependence on length can be eliminated by first sonicating the DNA to form fragments of about the same size. If the rates of renaturation are measured at the same DNA concentration in moles of nucleotides per liter, they vary by several orders of magnitude, as shown in Figure 6-13. The figure includes data for polyA–polyU and a double-stranded RNA (MS-2 viral RNA). The scale at the top is the number of nucleotide pairs in the genome. The larger the genome, the smaller the number of complementary fragments. Only complementary strands can form native DNA, so that the smaller the number of complementary strands, the slower the rate. For polyU and polyA, every strand is complementary so that this system is assigned a value of 1 on this scale. The renaturation process can be analyzed quantitatively. If the complementary strands are represented as A and A', the reaction can be written as



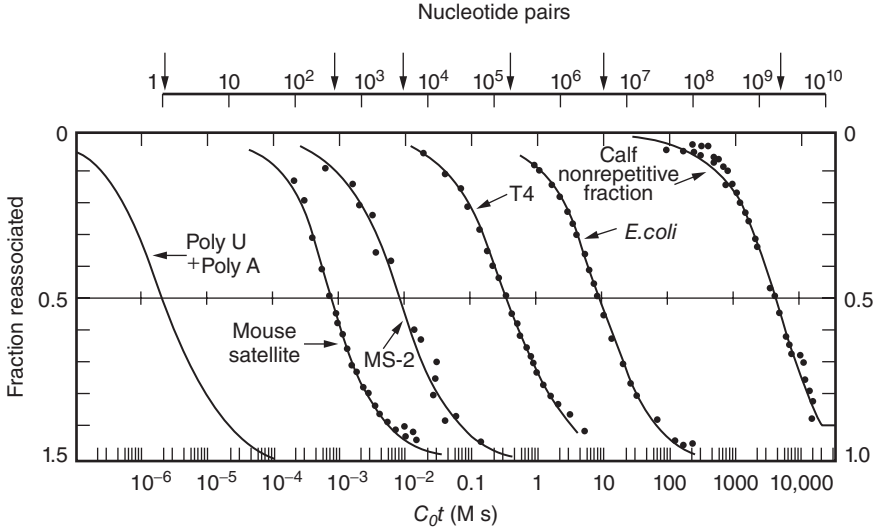


FIGURE 6-13. Reassociation of double-stranded nucleic acids from various sources. The genome size, which is a measure of the complexity of the DNA, is shown above the figure. The nucleic acids all have a single-stranded chain length of about 400 bases. Adapted with permission from R. J. Britten and D. E. Kohne, *Science* **161**, 529 (1968). Copyright © 1968 American Association for the Advancement of Science.

Since the concentrations of A and A' are equal, the rate equation is

$$-\frac{d[A]}{dt} = k[A]^2 \quad (6-45)$$

which can be integrated to give

$$\frac{1}{[A]} - \frac{1}{[A]_0} = kt \quad (6-46)$$

The data were found to obey this equation. The half-time for the reaction, that is, when $[A] = [A]_0/2$, is given by

$$t_{1/2} = \frac{1}{k[A]_0} \quad (6-47)$$

In the aforementioned equations, the concentration should be equal to the concentration of complementary strands, which we generally do not know. If the total number of base pairs in the smallest repeating sequence is N , then the concentration of complementary strands is proportional to the total concentration of DNA, c_0 , divided by N . If this relationship is substituted in Eq. (6-47), we find that $c_0 t_{1/2} \propto N$. Thus, the total concentration of DNA times the half-time for renaturation is a measure of N . A long half-time means that a given fragment will have to sample

many other fragments before it finds its complement, so the sequence is complex. N is a measure of the “complexity” of the DNA, and the scale on the top of Figure 6-13 can be equated to N . On this basis, the calf DNA is the most complex DNA that was studied. The study of the kinetics of renaturation can be a useful probe of a gene. It has been used, for example, to determine if fragments of a viral gene were present in viral transformed cells (20).

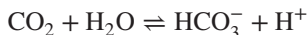
This concludes the sampling of the application of kinetics to biological systems. Many interesting studies have not been discussed, and many more remain to be done. The study of the time dependence of biological processes is essential in the quest for the elucidation of molecular mechanisms.

REFERENCES

1. A. Cornish-Bowden, *Fundamentals of Enzyme Catalysis* 4th edition, Portland Press, London, 2012
2. P. Frey and A. D. Hageman, *Enzymatic Reaction Mechanisms*, Oxford Press, New York, 2007.
3. D. L. Purich (ed.), *Contemporary Enzyme Kinetics and Mechanism* 3rd edition, Academic Press, New York, 2009.
4. B. S. Hartley and B. A. Kilby, *Biochem. J.* 56, 288 (1954).
5. G. P. Hess, in *The Enzymes*, 3rd edition, Vol. 3 (P. Boyer, ed.) Academic Press, New York, 1970 p. 213.
6. J. M. Denu, J. A. Stuckey, M. A. Saper, and J. E. Dixon, *Cell* 87, 361 (1996).
7. S. Shenolikar, *Annu. Rev. Cell Biol.* 10, 55 (1994).
8. J. M. Denu, D. L. Lohse, J. Vijayalakshmi, M. A. Saper, and J. E. Dixon, *Proc. Natl. Acad. Sci. USA* 93, 2493 (1996).
9. D. L. Lohse, J. M. Denu, N. Santoro, and J. E. Dixon, *Biochemistry* 36, 4568 (1997).
10. E. A. Doherty, *Annu. Rev. Biochem.* 69, 597 (2000).
11. E. A. Doherty and J. Doudna, *Annu. Rev. Biophys. Biomol. Structure* 30, 457 (2001).
12. D. M. Lilly, F. Eckstein, and S. Neidle (eds.) *Ribozymes and RNA Catalysis*, RSC Publishing, Cambridge UK, 2007.
13. A. Loria, S. Niranjankumari, C. A. Fierke, and T. Pan, *Biochemistry* 37, 15466 (1998).
14. J. A. Beebe and C. A. Fierke, *Biochemistry* 33, 10294 (1994).
15. G. G. Hammes and A. C. Park, *J. Am. Chem. Soc.* 91, 956 (1969).
16. T. G. Dewey and D. H. Turner, *Biochemistry* 18, 5757 (1979).
17. D. Riesner and R. Romer, in *Physico-Chemical Properties of Nucleic Acids*, Vol. 2 (J. Duchesne, ed.), Academic Press, London, 1973.
18. D. Pörschke and M. Eigen, *J. Mol. Biol.* 62, 361 (1971).
19. M. T. Record and B. H. Zimm, *Biopolymers* 11, 1435 (1972).
20. S. J. Flint, P. H. Gallimore, and P. A. Sharp, *J. Mol. Biol.* 96, 47 (1975).

PROBLEMS

- 6-1.** The hydration of CO_2 is catalyzed by the enzyme carbonic anhydrase. The overall reaction at neutral pH can be written as



The steady-state kinetics of both hydration and dehydration have been studied at pH 7.1, 0.5°C . Some typical data are given subsequently for an enzyme concentration of 2.8×10^{-9} M.

Hydration		Dehydration	
$10^3/v(\text{M}^{-1}\text{s})$	$10^3[\text{CO}_2](\text{M})$	$10^3/v(\text{M}^{-1}\text{s})$	$10^3[\text{CO}_3^-](\text{M})$
36	1.25	95	2.0
20	2.5	45	5.0
12	5.0	29	10
6.0	20	25	15

Calculate the steady-state parameters for the forward and reverse reactions.

- 6-2.** Studies of the inhibition of enzymes by various compounds often provide information about the nature of the binding site and the mechanism. Competitive inhibition is when the inhibitor, I, competes with the substrate for the catalytic site. This mechanism can be written as



Derive the steady-state rate law for this mechanism and show that it follows Michaelis–Menten kinetics when the inhibitor concentration is constant. Assume that the inhibitor concentration is much greater than the enzyme concentration.

- 6-3.** In the text, the steady-state rate law was derived with the assumption that the reaction is irreversible and/or only the initial velocity was determined. Derive the steady-state rate law for the reversible enzyme reaction:



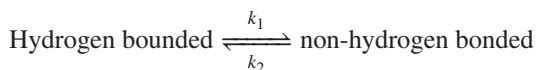
Show that the rate law can be put into the form

$$v = \frac{(V_S/K_S)[\text{S}] - (V_P/K_P)[\text{P}]}{1 + [\text{S}]/K_S + [\text{P}]/K_P}$$

where V_S and V_P are the maximum velocities for the forward and reverse reactions, and K_S and K_P are the Michaelis constants for the forward and reverse reactions.

When equilibrium is reached, $v = 0$. Calculate the ratio of the equilibrium concentrations of S and P, $[P]/[S]$, in terms of the four steady-state parameters. This relationship is called the Haldane relationship and is a method for determining the equilibrium constant of the overall reaction.

- 6-4.** The kinetics of the formation and breakdown of hydrogen-bonded loops or hairpins within small RNA molecules can be studied with relaxation methods since the formation of the helical structures is accompanied by a spectral change. This reaction can be represented as



What is the relaxation time for this reaction?

The relaxation time determined at a specific concentration was found to be $10 \mu\text{s}$. Calculate the individual rate constants if the equilibrium constant is 0.5.

If the concentration of the RNA is doubled, would the relaxation time get smaller, larger, or stay the same?

- 6-5.** Consider the binding of a protein, P, to a DNA segment (gene regulation). Assume that only one binding site for P exists on the DNA and that the concentration of DNA binding sites is much less than the concentration of P. The reaction mechanism for binding can be represented as



where the second step represents a conformational change in the protein. Calculate the rate law for the appearance of P-DNA' under the following conditions.

- The first step in the mechanism equilibrates rapidly relative to the rate of the overall reaction and $[\text{P-DNA}] \ll [\text{DNA}]$.
- The intermediate, P-DNA, is in a steady state.
- The first step in the mechanism equilibrates rapidly relative to the rate of the overall reaction and the concentrations of DNA and P-DNA are comparable. Express the rate law in terms of the *total* concentration of DNA and P-DNA, that is, $[\text{DNA}] + [\text{P-DNA}]$.
- The following initial rates were measured with an initial DNA concentration of $1 \mu\text{M}$.

[P] (μM)	10^4 rate (M/s)
100	8.33
50	7.14
20	5.00
10	3.33

Which of the rate laws is consistent with the data?

- 6-6.** Many biological reactions are very sensitive to pH. This can readily be incorporated into the rate laws because proteolytic reactions can be assumed to be much faster than other rates in most cases. For example, in enzyme mechanisms the ionization states of a few key protein side chains are often critical. Suppose that two ionizable groups on the enzyme are critical for catalytic activity and that one of them needs to be protonated and the other deprotonated. The proteolytic reactions can be written as



If only the species EH is catalytically active and the proteolytic reactions are much more rapid than the other steps in the reaction, all of the rate constants that multiply the free enzyme concentration in the rate law have to be multiplied by the fraction of enzyme present as EH.

- Calculate the fraction of free enzyme present as EH at a given pH. Your answer should contain the concentration of H^+ and the ionization constants of the two side chains, $K_{\text{E1}} = [\text{E}][\text{H}^+]/[\text{EH}]$ and $K_{\text{E2}} = [\text{EH}][\text{H}^+]/[\text{EH}_2]$.
- Assume that ES in the Michaelis–Menten mechanism (Eq. (6-2)) also exists in three protonation states, ESH_2 , ESH , and ES , with only ESH being catalytically active. Calculate the fraction of the enzyme–substrate complex present as EHS . Designate the ionization constants as K_{ES1} and K_{ES2} .
- Use the results of parts A and B to derive equations for the pH dependence of V_m , K_M , and V_m/K_M . Measurement of the pH dependence of the steady-state parameters permits determination of the ionization constants and sometimes identification of the amino acid side chains.

QUANTUM MECHANICS

Fundamentals of Quantum Mechanics

7.1 INTRODUCTION

Quantum mechanics was developed because of the failure of Newtonian mechanics to explain experimental results that emerged at the beginning of the 20th century. For example, for certain metals, such as Na, electrons are emitted when light is absorbed. This *photoelectric* effect has several nonclassical characteristics. First, for light of a given frequency, the kinetic energy of the electrons emitted is independent of the light intensity. The number of electrons produced is proportional to the light intensity, but all of the electrons have the same kinetic energy. Second, the kinetic energy of the photoelectron is zero until a threshold energy is reached, and then the kinetic energy becomes proportional to the frequency. This behavior is shown schematically in Figure 7-1, where the kinetic energy of the electrons is shown as a function of the frequency of the radiation. An explanation of these phenomena was proposed by Einstein, who, following Planck, postulated that energy is absorbed only in discrete amounts of energy, $h\nu$. A photon of energy $h\nu$ has the possibility of ejecting an electron, but a minimum energy is necessary. Therefore,

$$\text{kinetic energy} = h\nu - h\nu_0 \quad (7-1)$$

where h is Planck's constant, 6.625×10^{-27} erg-s (6.625×10^{-34} J-s) and $h\nu_0$ is the *work function* characteristic of the metal. This predicts that altering the light intensity would affect only the number of photoelectrons and not the kinetic energy. Furthermore, the slope of the experimental plot (Fig. 7-1) is Planck's constant, h .

On the other hand, many other experiments, including the interference patterns of light waves passing through two parallel slits, have illustrated the wave nature of light. Specifically, such experiments have shown that light can be described as a wave with a frequency, ν , or wavelength, λ , that is connected to its energy, E , by the familiar Planck relationship:

$$E = h\nu = \frac{hc}{\lambda} \quad (7-2)$$

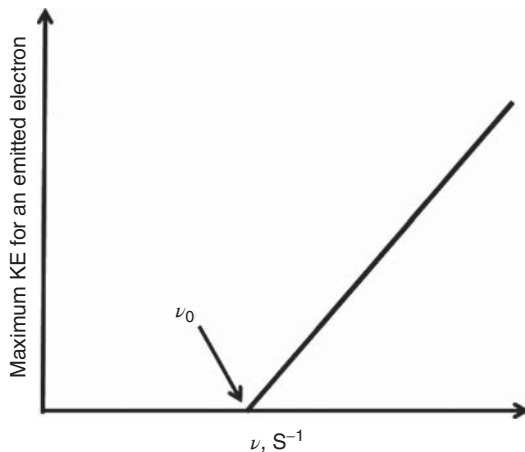


FIGURE 7-1. Schematic representation of the photoelectric effect. The maximum kinetic energy of an electron emitted from a metal surface when it is illuminated with light of frequency ν is shown. The frequency at which electrons are no longer emitted determines the work function, $h\nu_0$, and the slope of the line is Planck's constant, h (Eq. (7-1)).

where c is the speed of light, 2.998×10^{10} cm/s (2.998×10^8 m/s). Note that the product of the wavelength and the frequency is the speed of light: $\lambda\nu = c$. According to this description, light, or more generally radiation, can be envisaged as an electromagnetic sine wave that contains both electric and magnetic components, as shown in Figure 7-2. This figure shows that the electric component of the wave is perpendicular to the magnetic component and illustrates the relationship between the sine wave and the wavelength of the light. The physical properties of light will be discussed further in the context of spectroscopy in the last section of this chapter.

Returning to the photoelectric effect, the explanation given earlier postulates that light is corpuscular and consists of discrete photons characterized by a specific frequency or energy. How can this be reconciled with the well-known wave description of light? The answer is that both descriptions are correct – light can be envisioned as either discrete photons or a continuous wave. This wave-particle duality is a fundamental part of quantum mechanics. Although both descriptions are correct, one of them may more easily explain a given experimental situation.

About this point in history, de Broglie suggested that this duality is applicable to matter also, so that matter can be described as particles or waves. For light, the energy is equal to the momentum, p , times the speed of light, c , and by Einstein's postulate is also equal to $h\nu$:

$$E = h\nu = pc \quad (7-3)$$

Furthermore, since $\lambda\nu = c$, $p = h/\lambda$. For macroscopic objects, $p = mv$, where v is the velocity and m is the mass. In this case, $\lambda = h/(mv)$, the de Broglie wavelength. These fundamental relationships have been verified for matter by several experiments such as the diffraction of electrons by crystals.

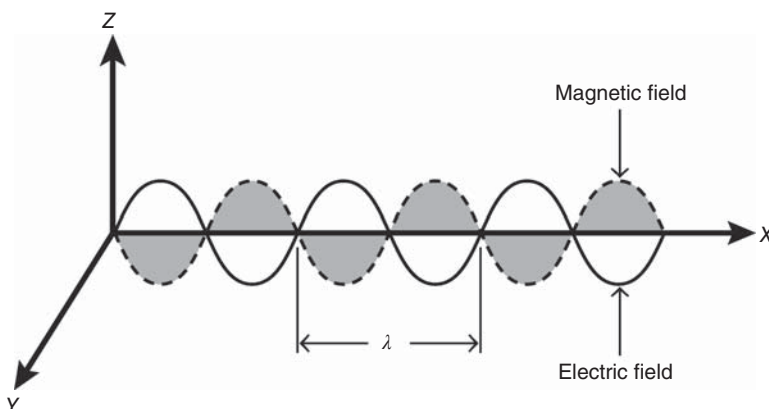


FIGURE 7-2. Schematic representation of an electromagnetic sine wave. The electric field is in the xz plane and the magnetic field in the xy plane. The electric and magnetic fields are perpendicular to each other at all times. The wavelength, λ , is the distance required for the wave to go through a complete cycle.

The postulate of de Broglie can be extended to derive an important result of quantum mechanics developed by Heisenberg in 1927, namely the uncertainty principle:

$$\Delta p \Delta x \geq \hbar \quad (7-4)$$

where $\hbar$ is defined to be Planck's constant h divided by 2π . In this equation, Δp represents the uncertainty in the momentum and Δx the uncertainty in the position. The uncertainty principle means that it is not possible to determine the precise values of the momentum, p , and the position, x . The more precisely one of these variables is known, the less precisely the other variable is known. This has no practical consequences for macroscopic systems but is crucial for the consideration of systems at the atomic level. For example, if a ball weighing 100 grams moves at a velocity of 100 miles/hour (a good tennis serve), an uncertainty of 1 mile/hour in the speed gives $\Delta p \sim 4.4 \times 10^{-22}$ kg m/s and $\Delta x \sim 2 \times 10^{-33}$ m. We are unlikely to worry about this uncertainty! On the other hand, if an electron (mass = 9×10^{-28} g) has an uncertainty in its velocity of 1×10^8 cm/s, the uncertainty in the position is about 1 Å, a large distance in terms of atomic dimensions. As we will see later, quantum mechanics has an alternative way of defining the position of an electron.

A second puzzling aspect of experimental physics in the late 1800s and early 1900s was found in the study of atomic spectra. Contrary to the predictions of classical mechanics, discrete lines at specific frequencies were observed when atomic gases at high temperatures emitted radiation. This can only be understood by the postulation of discrete energy levels for electrons. This was first explained by the famous Bohr atom, but this model was found to have shortcomings, and the final resolution of the problem occurred only when quantum mechanics was developed by Schrödinger and Heisenberg in the late 1920s. We will only consider the development by Schrödinger,

which is somewhat less complex than that of Heisenberg. The next section presents the famous Schrödinger equation, which is able to describe the previously puzzling results discussed earlier. For example, the wave-like behavior of electrons and the discrete energy levels of atoms are described via the Schrödinger equation. Other phenomena such as atoms tunneling through energy barriers can also be described with this formalism and will be discussed in the following sections.

7.2 SCHRÖDINGER EQUATION

Schrödinger postulated that all matter can be described as a wave and developed a differential equation that can be solved to determine the properties of a system. Basically, this differential equation contains two important contributions, the kinetic energy and the potential energy. Both of these are well-known concepts from classical mechanics, but they are redefined in the development of quantum mechanics. The one-dimensional Schrödinger equation for a particle of mass m has the form

$$-\frac{\hbar^2}{2m} \frac{d^2\psi_n(x)}{dx^2} + U(x)\psi_n(x) = E_n\psi_n(x) \quad (7-5)$$

In this equation, x is the position coordinate, and $U(x)$ is the potential energy of the system. For a given system with a known potential energy, this equation can be solved to obtain the wavefunctions, $\psi_n(x)$, and the associated energies, E_n , of the system. The subscript n refers to a specific quantum state of the system, and E_n is the energy of quantum state n . A more rigorous mathematical understanding of the Schrödinger equation requires the use of quantum mechanical operators and complex numbers and is beyond the scope of this book.

The Schrödinger equation given in Eq. (7-5) can be understood as the sum of kinetic energy and potential energy terms on the left side and a total energy term on the right side. Although these terms look very different from classical mechanics, the sum of the kinetic and potential energies is still the total energy in quantum mechanics. The key difference is that in classical mechanics, we know the exact location of the particle, whereas in quantum mechanics, we can only know the probability of finding the particle at any position x . The uncertainty principle discussed earlier (Eq. (7-4)) prevents us from knowing the exact location of a quantum mechanical particle, unless its momentum is completely unknown. However, the wavefunction in Eq. (7-5) gives us information about the probability of finding the particle at any point in space. In particular, $|\psi_n(x)|^2 dx$ is the probability of finding the particle between x and $x + dx$. For this reason, $|\psi_n(x)|^2$ is called the *probability density*. This probabilistic view of matter, denoted the *Born interpretation*, is in contrast to the deterministic character of Newtonian mechanics and has sparked considerable philosophical debate. In fact, Einstein apparently never fully accepted this probabilistic view of nature. (Albert Einstein: “God does not play dice with the universe.”)

It is important to realize that quantum mechanics reduces to classical mechanics for the heavy objects that we encounter in our everyday life. If a particle has a large mass, such as a tennis ball, its wavefunction will be highly localized around a certain

point in space, with a high probability of being at that position. Moreover, as discussed earlier, the uncertainties in the position and momentum of a macroscopic object have no practical consequences. Such heavy objects will behave classically and can be described by Newton's equations of motion, where force is the product of the mass and the acceleration. Lighter particles, such as electrons, are more delocalized and cannot be pinned down to a certain point in space. These light particles exhibit quantum mechanical behavior and do not obey Newton's equations of motion. In these cases, the average position of a particle can be calculated, but its specific position at a given time is unknown.

For the remainder of this chapter, we will present the solutions of the Schrödinger equation for different choices of the potential energy function associated with simple model systems. The results will illustrate fundamental quantum mechanical principles as well as concepts that are relevant to vibrations and rotations in molecules and to atomic and molecular spectra. The next chapter will extend these treatments to larger systems that are more directly relevant to biological systems.

7.3 PARTICLE IN A BOX

As an example of a simple quantum mechanical system that leads to quantization of energy levels, we consider a particle of mass m moving back and forth in a one-dimensional box of length L . This model actually has some practical application. It describes the movement of π electrons that are delocalized over a large part of a molecule, for example, biological molecules such as carotenoids, hemes, and chlorophyll. This is not an ordinary box because inside the box, the potential energy of the system is zero, whereas outside of the box, the potential energy is infinite, as depicted in Figure 7-3.

Because the potential walls are infinitely high, the solution to this equation outside of the box is easy—there is no chance the particle is outside the box, so the wave-function must be 0. Inside the box, $U = 0$, and Eq. (7-5) can be easily solved. The solution is

$$\psi_n(x) = A \sin bx \quad (7-6)$$

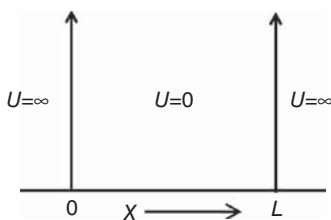


FIGURE 7-3. Quantum mechanical model for a particle in a one-dimensional box of length L . The particle is confined to the box by setting the potential energy equal to 0 inside the box and ∞ outside of the box.

where A and b are constants. At the ends of the box, $\psi_n(x)$ must be zero. This happens when $\sin n\pi = 0$ and n is an integer, so b must be equal to $n\pi/L$. This causes $\psi_n(x)$ to be 0 when $x = 0$ and $x = L$ for all integral values of n . To evaluate A , we use the concept that the probability of finding the particle in the interval between x and $x + dx$ is $|\psi_n(x)|^2 dx$. Because the particle must be in the box, the probability of finding the particle in the box is 1, so

$$\int_0^L |\psi_n(x)|^2 dx = \int_0^L A^2 \sin^2 \left(\frac{n\pi x}{L} \right) dx = 1 \quad (7-7)$$

Evaluation of this integral gives $A = \sqrt{2/L}$. Thus, the final result for the wavefunction is

$$\psi_n(x) = \sqrt{2/L} \sin \left(\frac{n\pi x}{L} \right) \quad (7-8)$$

Obviously n cannot be 0, as this would predict that there is no probability of finding the particle in the box, but n can be any nonzero positive integer. The wavefunctions for a few values of n are shown in Figure 7-4. As this figure illustrates, $\psi_n(x)$ is a sine wave, with the “wavelength” decreasing as n increases. Moreover, the number of *nodes* in the wavefunction, which is defined as the number of times the wavefunction changes sign, going from positive to negative or negative to positive, increases as n increases. The lowest state, with $n = 1$, has zero nodes, the next state, with $n = 2$, has one node, and the n th state has $n - 1$ nodes.

To determine the energy of the particle, all we have to do is put Eq. (7-8) back into Eq. (7-5) and solve for E_n . The resulting energy is

$$E_n = \frac{h^2 n^2}{8mL^2} \quad (7-9)$$

Thus, we see that the energy is quantized, and the energy is characterized by a series of energy levels, as depicted in Figure 7-5. Each energy level, E_n , is associated with

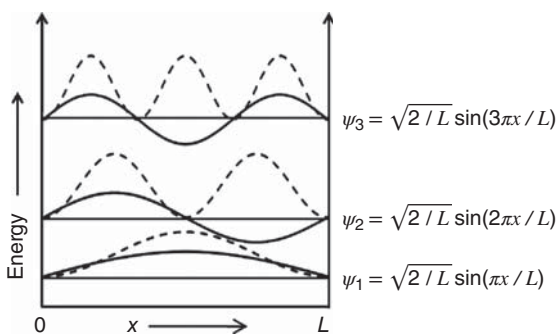


FIGURE 7-4. Wavefunctions, $\psi_n(x)$, for the first three energy levels of the particle in a box (Eq. (7-8)). The dashed lines show the probability of finding the particle at a given position x , as given by $|\psi_n(x)|^2$, scaled by a constant.

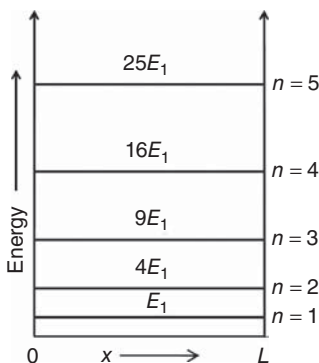


FIGURE 7-5. Energy levels for a particle in a box (Eq. (7-9)). The energy levels are expressed as n^2E_1 , where E_1 is the energy when $n = 1$.

a specific wavefunction, $\psi_n(x)$, and depends on the mass m of the particle and the length L of the box. Note that the energy levels would be very widely spaced for a very light particle, such as an electron, but would be very closely spaced for a macroscopic particle. Similarly, the smaller the box, the more widely spaced the energy levels. For a tennis ball being hit on a tennis court, the ball is sufficiently heavy and the court (box) sufficiently big so that the energy levels are a continuum for all practical purposes, and we do not even notice that the energy levels are quantized. The uncertainty in the momentum and position of the ball cannot be blamed on quantum mechanics in this case! The particle in a box illustrates how quantum mechanics can be used to calculate the properties of systems and how quantization of energy levels arises. The same calculation can be easily done for a three-dimensional box. In this case, the energies of the quantized states are the sum of three terms identical to Eq. (7-9), but with each of the three terms having a different quantum number.

The quantum mechanical description of matter does not permit the precise position of the particle to be determined, a manifestation of the Heisenberg uncertainty principle. However, the probability of finding the particle within a given segment of the box can be calculated. For example, the probability of finding the particle in the middle of the box, that is, between $L/4$ and $3L/4$ for the lowest energy state is

$$\int_{L/4}^{3L/4} \psi_1^2(x) dx = \left(\frac{2}{L}\right) \int_{L/4}^{3L/4} \sin^2\left(\frac{\pi x}{L}\right) dx \quad (7-10)$$

Evaluation of this integral gives a probability of 0.82. The probability of finding the particle within the middle part of the box is independent of L , the size of the box, but does depend on the value of the quantum number n . For the second energy level, $n = 2$, the probability is 0.50. The probability of finding the particle at position x in the box is shown as a dashed line for the first three energy levels in Figure 7-4. As n increases, the probability becomes more evenly distributed across the entire box, and for very high n , we reach the classical limit of equal probability everywhere in the

box, as one would envision if a tennis ball were bouncing back and forth between the two walls at constant speed.

An important result of quantum mechanics is that not only do molecules exist in different discrete energy levels, but the interaction of radiation with molecules causes shifts between these energy levels. If energy or radiation is *absorbed* by a molecule, the molecule can be raised to a higher energy state, whereas if a molecule loses energy, radiation can be *emitted*. For both cases, the change in energy is related to the radiation that is absorbed or emitted, namely, the change in energy state of the molecule, ΔE , is

$$\Delta E = h\nu = \frac{hc}{\lambda} \quad (7-11)$$

The change in energy, ΔE , is the difference in energy between specific energy levels of the molecule, for example, $E_2 - E_1$, where 1 and 2 designate different energy levels. It is important to note that since the energy is quantized, the light emitted or absorbed between two particular states is always a specific single frequency. Equation (7-11) can be applied to the particle in a box for the particle dropping from the $n + 1$ energy level to the n energy level:

$$\Delta E = \frac{h^2}{8mL^2}[(n+1)^2 - n^2] = \frac{h^2}{8mL^2}(2n+1) = \frac{hc}{\lambda} \quad (7-12)$$

If the particle is assumed to be an electron moving in a molecule 20 Å long and $n = 10$, then $\lambda \sim 600$ nm. This wavelength is in the visible region and has been observed for π electrons that are highly delocalized in molecules.

In practice, energy levels are sometimes so closely spaced that the frequencies of light emitted appear to create a continuum of frequencies. This is a shortcoming of the experimental method—in reality, the frequencies emitted are discrete entities. The particle in a box is a simple application of quantum mechanics, but it illustrates several important points that are also found in more complex calculations for molecular systems. First, the system can be described by a wavefunction. Second, this wavefunction permits determination of the probability of important characteristics of the system, such as positions. Finally, the energy of the system can be calculated and is found to be quantized. Moreover, the energy can only be absorbed or emitted in quantized packages characterized by specific frequencies. Quantum mechanical calculations also tell us what conditions are necessary for energy to be emitted or absorbed by a molecule. These calculations tell us *whether* radiation will be emitted or absorbed and what quantized packets of energy are available. These concepts will be developed further in the chapters focusing on spectroscopy.

7.4 VIBRATIONAL MOTIONS

The vibrational and rotational motions of molecules can also be described in terms of quantum mechanical models. A molecule comprised of N atoms has $3N$ coordinates, or degrees of freedom, because each atom can move in three-dimensional space. The

overall molecule can be described in terms of three translational motions associated with the entire molecule moving in three-dimensional space and two or three rotational motions for linear or nonlinear molecules, respectively. The remaining degrees of freedom are associated with vibrational motions: a linear molecule has $3N-5$ and a nonlinear molecule has $3N-6$ vibrational modes. For large macromolecules, the number of different vibrations can be approximated to be $3N$ because the number of translations and rotations is considered to be a small correction. In this section, we discuss the quantum mechanical description of the vibrational motions in molecular systems, including proteins and other large biomolecules.

A detailed analysis of all of the vibrational degrees of freedom of a macromolecule is not possible, but a few general conclusions bear mention. First, the characteristic vibration of each degree of freedom is a combination of the motions of many different bonds. These characteristic motions are called *normal modes* of vibration. The nature of these normal mode vibrations can be calculated very precisely for small molecules, and the characteristics of some of the normal modes can be associated with similar normal modes in macromolecules. Second, in some limiting cases, these normal modes are dominated by the movement of a single or a restricted number of chemical bonds. For example, the vibrations of C–H bonds, double bonds between C, and double bonds between C and O are each associated with similar spectral characteristics in many molecules. Third, the energy level distribution for each normal mode can be determined by approximating each normal mode as a simple harmonic oscillator, essentially a spring moving back and forth.

The remainder of this section will present the quantum mechanical description of a simple harmonic oscillator and will connect the resulting wavefunctions and energies to the vibrational motions in macromolecules. A simple harmonic oscillator is characterized by a potential energy of the form

$$U(x) = \frac{1}{2}kx^2 \quad (7-13)$$

In this equation, k is a constant and x is the vibrational coordinate. If the oscillator is thought of as a one-dimensional spring, the x coordinate is the coordinate along which the spring moves back and forth, and k is a measure of the stiffness of the spring. For a single bond, the x coordinate is the distance along the bond, and the origin of the coordinate system is the equilibrium bond distance. The characteristic frequency of the bond motion is

$$\nu = \left(\frac{1}{2\pi} \right) \left(\frac{k}{\mu} \right)^{1/2} \quad (7-14)$$

where μ is the reduced mass of the system. For a complex normal mode, the reduced mass is a weighted average of the masses; for our purposes, an exact calculation of the reduced mass is not necessary.

We can solve the Schrödinger equation for the simple harmonic oscillator by substituting the potential energy given in Eq. (7-13) into Eq. (7-5). The resulting wavefunctions, $\psi_n(x)$, have the form

$$\psi_n(x) = (\text{polynomial in } x) \times (\text{Gaussian function}) \quad (7-15)$$

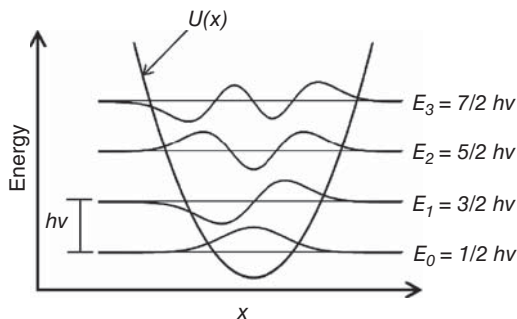


FIGURE 7-6. The potential energy curve and associated energy levels (Eq. (7-16)) and wavefunctions (Eq. (7-15)) for a simple harmonic oscillator. The energy levels are equally spaced with a separation of $h\nu$ between levels.

where the polynomial is a special function called a Hermite polynomial, and the Gaussian function is of the form $\exp\left(-\sqrt{mk/4\hbar^2}x^2\right)$. The energies associated with these wavefunctions are

$$E_n = \left(n + \frac{1}{2}\right) h\nu \quad (7-16)$$

in terms of the frequency ν of the vibrational motion. For this system, the quantum number n can be any nonnegative integer, including zero.

The potential energy for a harmonic oscillator, together with the energy levels and wavefunctions, are shown in Figure 7-6. Note that in the lowest energy state, $n = 0$, the normal mode still has an intrinsic energy, $h\nu/2$. This is called the zero-point energy and is possessed by all molecules, even at the hypothetical temperature of absolute zero. The zero-point energy for the particle in a box model is E_1 shown in Figure 7-5. The harmonic oscillator wavefunctions look qualitatively similar to the particle in a box wavefunctions, as can be seen by comparing Figures 7-4 and 7-6. In both cases, the effective “wavelength” decreases and the number of nodes increases as the quantum number n increases. On the other hand, there are several important differences between these two systems. The harmonic oscillator wavefunctions have nonzero probability beyond the edge of the potential energy curve into what is called the classically forbidden region, as will be discussed in the next section. This was not true for the particle in a box because the energy barriers between inside and outside the box were assumed to be infinite. Additionally, the energy levels are equally spaced with constant splittings of $h\nu$ for the harmonic oscillator (Figure 7-6), in contrast to the increasing splittings for the particle in a box (Figure 7-5).

For real molecules, the vibrational motion deviates from this simple harmonic model at high energies, and the energy levels become more closely spaced at the top of the potential well. The approximate potential energy function for a vibrational coordinate of a real molecule (anharmonic oscillator) and the associated energy levels are depicted in Figure 7-7. For a diatomic molecule, the x coordinate is the internuclear distance, whereas for larger molecules it is a combination of internuclear distances associated with the normal mode vibration. Quantum mechanical calculations can be

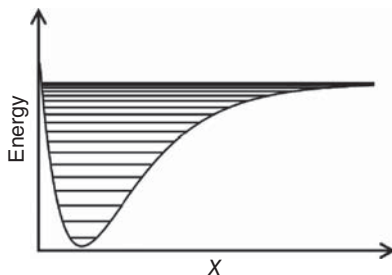


FIGURE 7-7. Energy levels for an anharmonic oscillator. The energy levels become more closely spaced at higher energies.

carried out with anharmonic potential energy functions to provide direct correlation of theory with experimental findings. Note that near the bottom of the potential well, harmonic and anharmonic oscillators behave quite similarly so that the harmonic oscillator is a good model at room temperature, where only the lowest-energy vibrational levels are accessible.

For a C–H bond, the reduced mass can be approximated as the mass of the hydrogen atom. A useful tool for identifying the nature of a given vibrational mode is to substitute deuterium for hydrogen. According to Eq. (7-14), the ratio of the characteristic frequency of the vibration is

$$\frac{\nu_{\text{H}}}{\nu_{\text{D}}} = \left(\frac{m_{\text{D}}}{m_{\text{H}}} \right)^{1/2} = \sqrt{2} \quad (7-17)$$

This substitution of deuterium for hydrogen is very useful for determining if the motion of a hydrogen atom is the dominant factor in a given vibrational mode. The same procedure works for an N–H bond, an O–H bond, and so forth.

7.5 TUNNELING

Tunneling is the phenomenon in which a particle penetrates into or through a region that is classically forbidden. If the particle were treated classically, it would not be able to enter the classically forbidden region. For example, if we throw a tennis ball at a solid wall, it will bounce back to us because it behaves classically. If the tennis ball behaved quantum mechanically, however, there would be some probability that it would tunnel through the wall and end up on the other side. The large mass of the ball prevents this type of quantum mechanical tunneling, but electrons and sometimes even hydrogen atoms are light enough that they are able to penetrate barriers in this way.

In classical mechanics, the total energy is defined to be the sum of the kinetic energy and the potential energy, and the kinetic energy is always positive. Therefore, the total energy can never be less than the potential energy in classical mechanics.

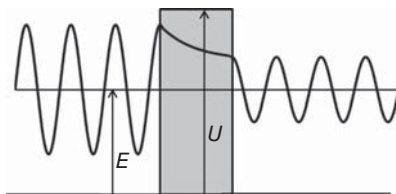


FIGURE 7-8. Schematic illustration of a particle tunneling through a potential energy barrier of height U , which is greater than the total energy, E , of the particle. The potential energy is zero outside the shaded region and U inside the shaded region. This figure illustrates tunneling into a classically forbidden region as the wavefunction propagates from left to right. For simplicity, the reflection is not shown.

In quantum mechanics, however, this situation is possible. An example is depicted in Figure 7-8, where the total energy E becomes less than the potential energy U at a wall that has a finite height and width. In classical mechanics, if a particle on the left side of the wall moves toward this wall with energy $E < U$, it would hit the wall and bounce back again, remaining on the left side of the wall because it does not have enough energy to surmount the wall. In quantum mechanics, if a particle is in a quantum state with energy $E < U$, then it could still get through the wall and into the region on the other side. This possibility can be understood by solving the Schrödinger equation for this system. As depicted in Figure 7-8, the particle is oscillating on the left side of the wall with a wavefunction that resembles a particle in a box wavefunction. However, it also has a nonzero probability of being in the region of the wall itself, where it decays smoothly, as well as in the region on the right side of the wall, where it starts to oscillate again.

Another example of tunneling is seen in the harmonic oscillator wavefunctions from the previous section. At the edges of the potential energy curve, the wavefunctions penetrate into the classically forbidden region, as shown in Figure 7-6. In these regions, the energy of the quantum state is less than the potential energy, yet the particle still has finite probability of occupying these regions. The classical analog can be understood by envisioning a ball that is rolling back and forth on a curved (harmonic) track. When the ball has zero kinetic energy at the top of its trajectory, it stops and turns around and rolls back in the opposite direction. By contrast, a quantum mechanical particle can go beyond this point and penetrate into the classically forbidden region.

A biologically relevant example of tunneling is hydrogen transfer in proteins. In these situations, the hydrogen nucleus can tunnel from one heavy atom to another. At a protein configuration that is conducive to tunneling, the hydrogen nucleus can be viewed as moving in a symmetric double-well potential, as shown in Figure 7-9. The solution of the Schrödinger equation for this system results in the wavefunctions depicted in this figure. These wavefunctions illustrate that the hydrogen nucleus has finite probability of penetrating the barrier region, which is classically forbidden. The splitting between the lowest two hydrogen vibrational energy levels is denoted the tunneling splitting and determines the probability of tunneling. An experimentally

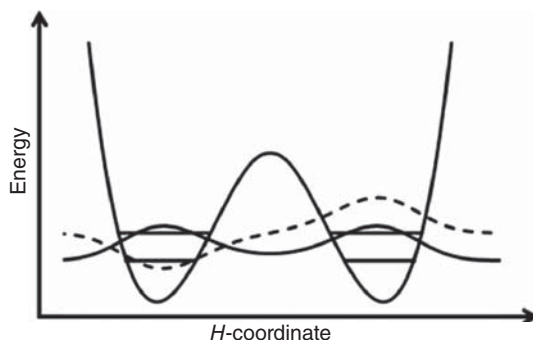


FIGURE 7-9. Schematic depiction of the potential energy curve along the hydrogen coordinate for a hydrogen transfer reaction and the lowest two energy levels and wavefunctions associated with the motion of the hydrogen nucleus. The solid and dashed lines show the wavefunctions for the first and second energy levels, respectively, and the tunneling splitting (enhanced for visualization purposes) is the difference between these two energy levels. The penetration of the wavefunctions into the classically forbidden barrier region illustrates hydrogen tunneling through the barrier.

measurable quantity that reflects hydrogen tunneling is the kinetic isotope effect, defined as the ratio of the rate constants for hydrogen and deuterium transfer. The tunneling probabilities for hydrogen and deuterium transfer differ because the wavefunctions depend strongly on the mass of the particle: deuterium is heavier and therefore acts more classically and is more localized in space, with smaller splittings between energy levels and a lower probability of tunneling. Thus, the kinetic isotope effect can become very large when tunneling is occurring. For example, the net hydrogen atom transfer catalyzed by the enzyme soybean lipoxygenase exhibits a kinetic isotope effect of ~ 80 at room temperature. This unusually high kinetic isotope effect can only be explained through quantum mechanical tunneling. Note that tunneling of other atoms that are heavier than hydrogen and deuterium is possible but much less likely because the tunneling probability depends exponentially on the mass.

7.6 ROTATIONAL MOTIONS

The rotational motions of macromolecules do not tend to play an important biological role, but the form of the rotational wavefunctions is a significant component of the hydrogenic atom wavefunctions, which in turn are used as the foundation to describe the electronic structure of larger biomolecules. In this section, we briefly discuss the quantum mechanical description of the rotational motions of a molecule. For the simplest case of a diatomic molecule, the system can be described as a rotating dumbbell with one atom at each end of a massless stick of length r_0 , the bond length of the molecule. This system can be converted to an even simpler model in terms of the center of mass and the reduced mass μ associated with the two atoms. In this simpler

model, a particle with mass μ is on one end of a stick of length r_0 rotating freely in space with the other end held fixed. Alternatively, this system can be described in terms of a particle moving on the surface of a sphere, envisioned as a small rubber ball moving on the surface of a much larger beach ball with radius r_0 . These two models are completely equivalent and require the introduction of spherical coordinates with radial coordinate r and two angular coordinates θ and φ . The Schrödinger equation is more complex than Eq. (7-5) because three dimensions are involved. The potential energy is zero, so the left side of the Schrödinger equation includes only a kinetic energy term.

In this model, the radial coordinate is fixed at r_0 , and the wavefunctions depend only on the two angular coordinates. The wavefunctions are special functions called spherical harmonics, typically denoted $Y_{l,m_l}(\theta, \varphi)$, and depend on two quantum numbers, l and m_l . The quantum numbers l and m_l reflect the magnitude and direction, respectively, of the angular momentum of the particle. The magnitude of the total angular momentum is $[l(l+1)]^{1/2}\hbar$, and the z -component of this angular momentum is $m_l\hbar$. For this reason, l is called the angular momentum quantum number. The energies of the quantum states depend on only the l quantum number and have the form

$$E_l = \frac{l(l+1)\hbar^2}{2mr_0^2} \quad (7-18)$$

The angular momentum quantum number l can be any nonnegative integer, including zero, and the other quantum number m_l can be any integer between $-l$ and l , leading to $2l+1$ quantum states with the same energy E_l . Quantum states that have the same energy are called degenerate, and the degeneracy of the level is the number of states with the same energy. Similar to the particle in a box, the spacing between the rotational energy levels increases as the quantum number l increases.

This model gives the rotational states that would be applicable to a diatomic molecule with fixed bond length. When it is extended to more complicated molecules, the energy levels have the same form, but the quantum number is often denoted J instead of l , and the denominator of the energy in Eq. (7-18) becomes $2I$, where I is the moment of inertia of the molecule. Moreover, often the energy levels are written in terms of a rotational constant defined as $B = \hbar/(4\pi I)$ to facilitate the analysis of rotational spectra:

$$E_J = hBJ(J+1) \quad (7-19)$$

Typically the splittings between the rotational energy levels are much smaller than the splittings between the vibrational energy levels. As a result, at room temperature, many of the rotational energy levels are populated, whereas only the lowest vibrational energy level is populated. As mentioned earlier, the rotational states will be used in the description of hydrogenic atoms in Chapter 8. Before moving on to the electronic structure of atoms and molecules in Chapter 8, however, we will summarize the basic concepts of spectroscopy to provide the foundation for future chapters that will focus on specific types of spectroscopy.

7.7 BASICS OF SPECTROSCOPY

Spectroscopy is a powerful tool for studying biological systems. It often provides a convenient method for the analysis of individual components in a biological system such as proteins, nucleic acids, and metabolites. It can also provide detailed information about the structure and mechanism of biomolecules. In order to obtain the maximum benefit from this tool and to use it properly, a basic understanding of the fundamentals of spectroscopic phenomena, as well as of the instrumentation currently available, is necessary. A detailed understanding involves complex theory, but a grasp of the important concepts and their applications can be obtained without resorting to advanced mathematics and theory. We will attempt to do this by emphasizing the physical ideas associated with spectral phenomena and utilizing a few of the concepts and results from molecular theory.

Very simply stated, spectroscopy is the study of the interaction of radiation with matter. Radiation is characterized by its energy, E , which is linked to the frequency, ν , and wavelength, λ , of the radiation by the Planck relationship given in Eq. (7-2). The useful wavelength of radiation for spectroscopy extends from X-rays, $\lambda \sim 1 - 100\text{ nm}$, to microwaves, $\lambda \sim 10^5 - 10^6\text{ nm}$. For biology, the most useful radiation for spectroscopy is in the ultraviolet and visible region of the spectrum. The entire useful spectrum is shown in Figure 7-10, along with the common names for the various regions of the spectrum. If radiation is envisaged as both an electric and a magnetic wave, then its interactions with matter can be considered as electromagnetic phenomena because matter is made up of positive and negative charges. We will not be concerned with the details of this interaction but will focus on the aspects that are essential to spectroscopic methods.

It is useful to consider the fundamental properties of light waves to better understand some of the experimental methods that will be discussed later. Thus far, we have considered light to be a periodic electromagnetic wave in space that could be characterized, for example, by a sine function:

$$I = I_0 \sin\left(\frac{2\pi x}{\lambda}\right) \quad (7-20)$$

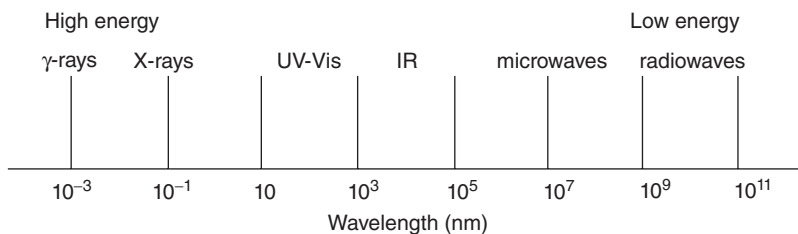


FIGURE 7-10. Schematic representation of the wavelengths associated with electromagnetic radiation. The wavelengths, in nanometers, span 14 orders of magnitude. The common names of the various regions are indicated approximately (UV is ultraviolet, Vis is visible, and IR is infrared).

Here I is the electric or magnetic field, I_0 is the maximum value of the electric or magnetic field, x is the distance along the x -axis, and λ is the wavelength. A light wave can also be periodic in time, as illustrated in Figure 7-11. In this case,

$$I = I_0 \sin(2\pi\nu t) = I_0 \sin(\omega t) \quad (7-21)$$

Now, I is the light intensity, I_0 is the maximum light intensity, ν is the frequency in s^{-1} , as defined in Figure 7-11, ω is the frequency in radians ($\omega = 2\pi\nu$), and t is the time. The velocity of the propagating wave is $\lambda\nu$, which in the case of electromagnetic radiation is the speed of light, that is, $\lambda\nu = c$. If light of the same frequency and maximum amplitude from two sources is combined, the two sine functions will be added. If the two light waves start with zero intensity at the same time ($t = 0$), the two waves add and the intensity is doubled. This is called *constructive interference*. If the two waves are combined with one of the waves starting at zero intensity and proceeding to positive values of the sine function, whereas the other begins at zero intensity and proceeds to negative values, the two intensities cancel each other out. This is called *destructive interference*. Obviously, it is possible to have cases in between these two extremes. In such cases, a phase difference is said to exist between the two waves. Mathematically, this can be represented as

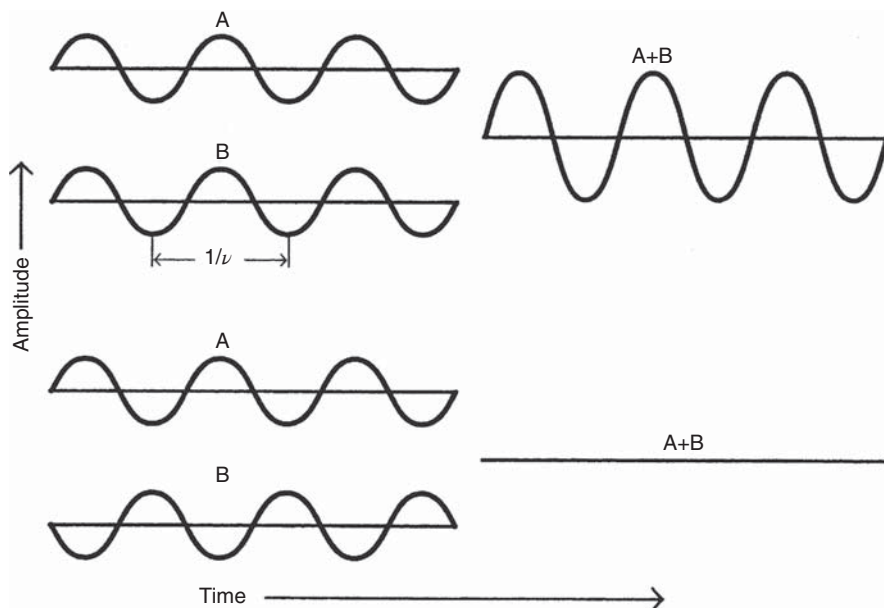


FIGURE 7-11. Examples of constructive and destructive interference. Constructive interference: when the upper two wave forms of equal amplitude and a phase angle of 0° (or integral multiples of 2π) are added (left), a sine wave with twice the amplitude and the same frequency results (right). Destructive interference: when the lower two wave forms are added (left), the amplitudes of the two waves cancel (right). The phase angle in this case is 90° (or odd integral multiples of $\pi/2$).

$$I = I_0 \sin(\omega t + \delta) \quad (7-22)$$

where δ is called a phase angle and can be either positive or negative. When many different waves of the same frequency are combined, the intensity can always be described by such a relationship. These phenomena are shown schematically in Figure 7-11.

A standard way of carrying out spectroscopy is to apply continuous radiation and then look at the intensity of the radiation after it has passed through the sample of interest. The intensity is then determined as a function of the frequency of the radiation, and the result is the absorption spectrum of the sample. The color of a material is determined by the wavelength of the light absorbed. For example, if white light shines on blood, blue/green light is absorbed so that the transmitted light is red. Several examples of absorption spectra are shown in Figure 7-12. We will consider why and how much the sample absorbs light a bit later, but you are undoubtedly already familiar with the concept of an absorption spectrum.

The use of continuous radiation is an effective way to carry out an experiment, but there is an interesting mathematical relationship that permits a different approach to the problem. This mathematical operation is the *Fourier transform*. The principle of a Fourier transform is that if the frequency dependence of the intensity, $I(\nu)$, can be determined, it can be transformed into a new function, $F(t)$, that is a function of the time, t . Conversely, $F(t)$ can also be converted to $I(\nu)$. Both of these functions contain the same information. Why then are these transformations advantageous? It can be quite time consuming to determine $I(\nu)$, but a short pulse of radiation can be applied very quickly. Basically, what this transformation means is that looking at the response of the system to application of a pulse of radiation, such as that shown in Figure 7-13, is equivalent to looking at the response of the system to sine wave radiation at many different frequencies. In other words, a square wave is mathematically equivalent to adding up many sine waves of different frequencies, and vice versa. This is shown schematically in Figure 7-13, where the addition of sine waves with four different frequencies produces a periodic “square” wave. The larger the number of sine waves added, the more “square” the wave becomes. In mathematical terms, a square wave can be represented as an infinite series of sine functions, a Fourier series.

The mathematical equivalence of timed pulses and continuous waves of many different frequencies has profound consequences in determining the spectroscopic properties of materials. In many cases, the use of pulses permits thousands of experiments to be done in a very short time. The results of these experiments can then be averaged, producing a far superior frequency spectrum in a much shorter time than could be determined by continuous wave methods. In later chapters, we will be dealing with both continuous wave spectroscopy and Fourier transform spectroscopy. It is important to remember that both methods give identical results. The method of choice is the one that produces the best data in the shortest time, and in some cases at the lowest cost.

With this brief introduction to the underlying theoretical principles of spectroscopy, we are ready to proceed with consideration of specific types of spectroscopy and their application to biological systems. Before proceeding to these chapters on

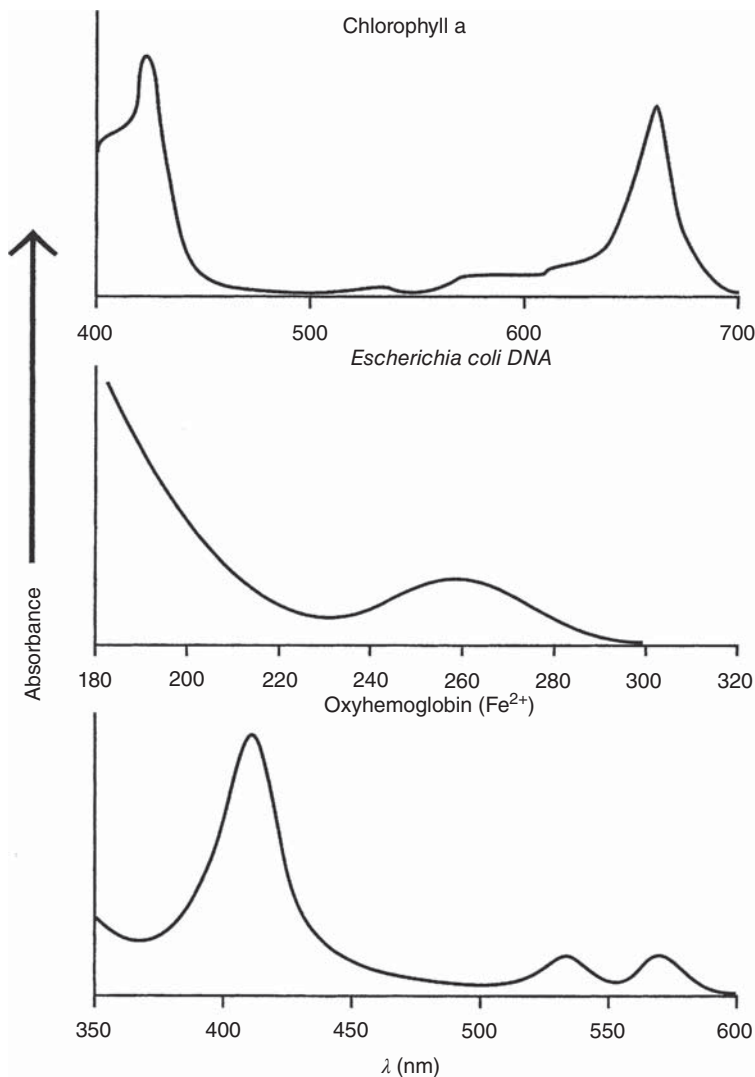


FIGURE 7-12. Absorption of light by biological molecules. The absorbance scale is arbitrary and the wavelength, λ , is in nanometers. Chlorophyll *a* solutions absorb blue and red light and are green in color. DNA solutions absorb light in the ultraviolet and are colorless. Oxyhemoglobin solutions absorb blue light and are red in color.

spectroscopy, however, we discuss the electronic structure of atoms and molecules in the next chapter. Understanding the electronic structure of atoms and molecules is important for interpreting the results of spectroscopic studies. Moreover, the application of electronic structure methods to biomolecules provides structural, energetic, and mechanistic insight into biological systems.

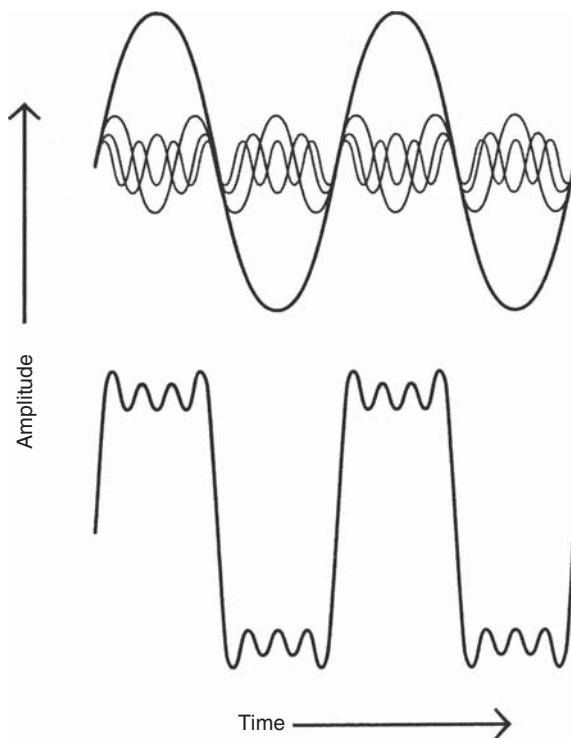


FIGURE 7-13. The upper part of the figure shows sine waves of four different frequencies, and the lower part of the figure is the sum of the sine waves, which approximates a square wave pulse of radiation. When sine waves of many more frequencies are included, the time dependence becomes a pulsed square wave. This figure illustrates that the superposition of multiple sine waves is equivalent to a square wave pulse, and vice versa. This equivalency is the essence of Fourier transform methods. Copyright by Professor T. G. Oas, Duke University. Reproduced with permission.

REFERENCES

The topics in this chapter are discussed in considerably more depth in a number of physical chemistry textbooks, such as those cited below (cf. Refs. 1–5).

1. I. Tinoco, Jr., K. Sauer, & J. C. Wang, J. D. Puglisi, G. Harbison, and D. Rovnyak, *Physical Chemistry: Principles and Applications to the Biological Sciences*, 5th edition, Prentice Hall, Englewood Cliffs, NJ, 2013.
2. R. J. Silbey, R. A. Alberty, and M. G. Bawendi, *Physical Chemistry*, 4th edition, John Wiley & Sons, New York, 2004.
3. P. Atkins and J. De Paula, *Physical Chemistry*, 9th edition, W. H. Freeman and Co., New York, NY, 2010.
4. R. S. Berry, S. A. Rice, and J. Ross, *Physical Chemistry*, 2nd edition, Oxford University Press, New York, 2000.

5. D. A. McQuarrie and J. D. Simon, *Physical Chemistry: A Molecular Approach*, University Science Books, Sausalito, CA, 1997.

PROBLEMS

- 7-1. The energies required to break the C–C bond in ethane, the “triple bond” in CO, and a hydrogen bond are about 88, 257, and 4 kcal/mol, respectively. What wavelengths of radiation are required to break these bonds?
- 7-2. Calculate the energy and momentum of a photon with the following wavelengths: 150 pm (X-ray), 250 nm (ultraviolet), 500 nm (visible), and 1 cm (microwave).
- 7-3. The maximum kinetic energy of electrons emitted from Na at different wavelengths was measured with the following results.

λ (Å)	Maximum kinetic energy (eV)
4500	0.40
4000	0.76
3500	1.20
3000	1.79

Calculate Planck’s constant and the value of the work function from these data. ($1 \text{ eV} = 1.602 \times 10^{-19} \text{ J}$.)

- 7-4. Calculate the de Broglie wavelength for the following cases:
- An electron in an electron microscope accelerated with a potential of 100 kV.
 - A He atom moving at a speed of 1000 m/s.
 - A bullet weighing 1 g moving at a speed of 100 m/s.
 - Assume that the uncertainty in the speed is 10% and calculate the uncertainty in the position for each of the three cases.
- 7-5. The particle in a box is a useful model for electrons that can move relatively freely within a bonding system such as π electrons. Assume that an electron is moving in a “box” that is 50 Å long, that is, a potential well with infinitely high walls at the boundaries.
- Calculate the energy levels for $n = 1, 2$, and 3.
 - What is the wavelength of light emitted when the electron moves from the energy level with $n = 2$ to the energy level with $n = 1$?
 - What is the probability of finding the electron between 12.5 and 37.5 Å for $n = 1$?
- 7-6. The zero-point energies for two different vibrational manifolds are 9.55×10^3 and $1.19 \times 10^4 \text{ J/mol}$.

- a. Calculate the frequency of light emitted for a transition between the first energy level and the zero-point energy for each of these manifolds.
 - b. In terms of the harmonic oscillator model, which of these manifolds has the “stiffer” spring?
 - c. For both cases, the radiation associated with a transition between energy levels 20 and 19 occurs at a longer wavelength (smaller wave number) than the radiation associated with a transition between the energy levels 2 and 1. How do you explain this?
- 7-7.** Sketch the graph of I versus t for sine wave radiation that obeys the relationship $I = I_0 \sin(\omega t + \delta)$ for $\delta = 0, \pi/4, \pi/2$ and π . Plot the sum of the sine waves when the sine wave for $\delta = 0$ is added to that for $\delta = 0$ or $\pi/4$, or $\pi/2$, or π . This exercise should provide you with a good understanding of constructive and destructive interference. Do your results depend on the value of ω ? Briefly discuss what happens when waves of different frequency are added together.

Electronic Structure of Atoms and Molecules

8.1 INTRODUCTION

In Chapter 7, we introduced the fundamental concepts of quantum mechanics and investigated several model systems that could be solved exactly. We started with the particle in a box, which can be used as a model for describing π -bonding systems, and then we discussed models for describing the vibrational and rotational motions of molecules. In particular, the harmonic oscillator model is very useful for describing the vibrational modes of macromolecules.

In this chapter, we investigate the electronic structure of atoms and molecules. First we discuss the electronic wavefunctions for hydrogenic atoms, which are comprised of one nucleus and one electron. We are able to solve this problem analytically with exact mathematical methods. Because large macromolecules contain many nuclei and many electrons, however, the quantum mechanical description of such systems becomes more complicated. Thus, we need to make several approximations to be able to study large macromolecules with quantum mechanical methods, and the problems can no longer be solved analytically. Instead, we need to develop computer algorithms to solve the Schrödinger equation numerically for larger systems. In the later sections of this chapter, we describe the approximations that must be made and discuss the computational methods that have been developed to allow the quantum mechanical study of biomolecules. We also present several examples of how quantum chemistry can provide insights into biological systems.

8.2 HYDROGENIC ATOMS

In Chapter 7, we focused mainly on the nuclear motions, specifically vibrations and rotations. However, molecules are made up of atoms, which in turn are made up of both nuclei and electrons. A complete understanding of molecular systems requires us to consider the motion of the electrons as well as the nuclei. In this section, we focus on hydrogenic atoms, which are comprised of a nucleus with charge Ze and an

electron with charge $-e$. Here Z is the atomic number for a neutral atom. The potential energy of this system is $U(r) = -Ze^2/r$, where r is the distance between the electron and the nucleus. For this system, we can solve the Schrödinger equation exactly, but as soon as we have more than a single electron, we need to make significant approximations. These more complicated systems, as well as applications to biomolecules, will be discussed later in this chapter.

The solution of the Schrödinger equation for hydrogenic atoms requires us to separate the system into a radial part and an angular part. In this case, such a separation is rigorous because the potential energy has no angular component. The wavefunctions for the angular part are the same as those obtained for the particle on a sphere discussed in Chapter 7 because we can view the nucleus as fixed in space, and the electron is moving on a sphere of radius r centered at the nucleus for each value of r . Thus, as explained in Chapter 7, the angular part of the wavefunction is given by the wavefunctions $Y_{l,m_l}(\theta, \varphi)$ and is described in terms of the same two quantum numbers l and m_l . The radial part of the hydrogenic wavefunction depends on the principal quantum number n and the angular momentum quantum number l and is of the form

$$R_{nl}(r) = r^l \times (\text{polynomial in } r) \times (\text{decaying exponential in } r) \quad (8-1)$$

The total wavefunction is of the form $Y_{l,m_l}(\theta, \varphi)R_{nl}(r)$, where the principal quantum number n can be any positive integer, the angular quantum number l can be any non-negative integer up to $n-1$, and m_l can be any integer from $-l$ to l .

Describing each wavefunction with three quantum numbers may seem complicated, but each quantum number has a physical meaning. The principal quantum number n determines the energy of the state. The energies of the quantum states are

$$E_n = -Z^2/n^2 \quad (8-2)$$

in units of Rydberg, given as $R_H = 2.180 \times 10^{-18}$ J. As shown in Figure 8-1, the splittings between the energy levels decrease as the quantum number n increases. Eventually the energy levels become continuous, and the electron separates completely from the nucleus, a process called ionization. The quantum numbers l and m_l designate the magnitude and direction, respectively, of the angular momentum associated with the electron moving around the nucleus. Specifically, the angular momentum quantum number l determines the magnitude of the total angular momentum, $[l(l+1)]^{1/2}\hbar$, and the third quantum number m_l determines the z -component of the angular momentum, $m_l\hbar$. These three quantum numbers also dictate the shapes and orientations of the wavefunctions.

Analysis of these wavefunctions leads to the familiar hydrogenic orbitals, where s , p , and d correspond to the quantum numbers $l=0$, 1 , and 2 , respectively. Here an orbital is simply a one-electron wavefunction. The hydrogenic orbitals are typically denoted as follows: $1s$ ($n=1, l=0$), $2s$ ($n=2, l=0$), $2p$ ($n=2, l=1$), $3s$ ($n=3, l=0$), $3p$ ($n=3, l=1$), $3d$ ($n=3, l=2$), and so forth. As for the model used to describe molecular rotations discussed in Chapter 7, the degeneracy of each energy level with quantum number l is $2l+1$. This degeneracy for $l=1$ leads to the three $2p_x$, $2p_y$, and

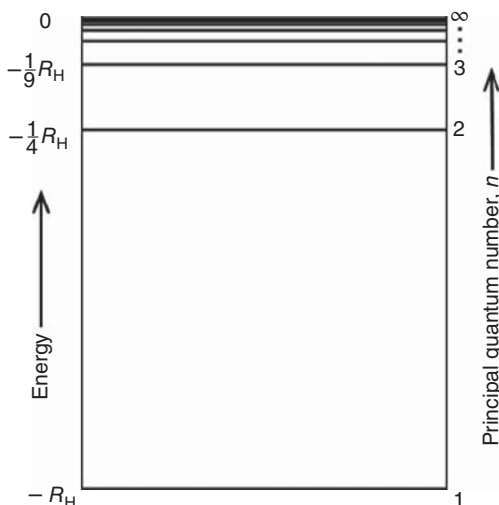


FIGURE 8-1. Energies of the hydrogen atom orbitals, illustrating that the splittings between energy levels become smaller as the principal quantum number n increases. Here R_H denotes the Rydberg unit of energy.

TABLE 8-1. Lowest Five Hydrogen Atom Orbitals^a

Orbital	n	l	Wavefunction	Energy
$1s$	1	0	$\exp(-\rho/2)$	$-1 R_H$
$2s$	2	0	$(2 - \rho) \exp(-\rho/2)$	$-0.25 R_H$
$2p_x$	2	1	$x \exp(-\rho/2)$	$-0.25 R_H$
$2p_y$	2	1	$y \exp(-\rho/2)$	$-0.25 R_H$
$2p_z$	2	1	$z \exp(-\rho/2)$	$-0.25 R_H$

^a n and l are the principal and angular momentum quantum numbers, $\rho = 2r/na_0$, $r = (x^2 + y^2 + z^2)^{1/2}$ in Cartesian coordinates, a_0 is the Bohr radius with a value of 52.9 pm, and R_H denotes the Rydberg unit of energy. Each wavefunction should also be multiplied by a constant to ensure that the total probability over all space is unity.

$2p_z$ orbitals that are aligned along the x , y , and z axes, respectively. For hydrogenic atoms, the energies for all states with quantum number n are degenerate (Eq. (8-2)). Thus, the degeneracy of the $n = 2$ level is four (one $2s$ and three $2p$ orbitals).

The wavefunctions and energies of the $1s$, $2s$, and $2p$ orbitals are given in Table 8-1, and the wavefunctions are depicted in Figure 8-2. The $1s$ orbital has no nodes and has a maximum value at the nucleus, decaying exponentially as a function of the radial coordinate r . The $2s$ orbital looks similar but has one radial node, which can be envisioned as a ball of radius $\rho = 2$, where the wavefunction is zero on the surface of the ball and has a different sign inside the ball than it has outside the ball. The $2p$ orbitals do not have a radial node but have an angular node, which means

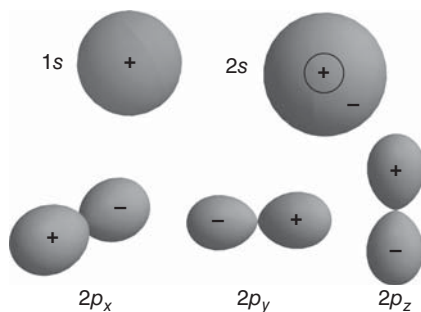


FIGURE 8-2. Wavefunctions associated with the $1s$, $2s$, $2p_x$, $2p_y$, and $2p_z$ hydrogen atomic orbitals. The plus and minus signs indicate the sign of the wavefunction. The $2s$ orbital has one radial node and the $2p$ orbitals each have one angular node.

that the wavefunction is zero on an entire plane, called a nodal plane, and changes sign across this plane. For example, the $2p_x$ orbital is zero in the yz plane and has a different sign for positive and negative values of x . These general trends continue for the higher energy states: the number of radial nodes is $n - l - 1$, and the number of angular nodes is l . For example, the $3d$ orbitals have no radial nodes and two angular nodes and therefore have more complicated shapes that arise from the two nodal planes.

We can also interpret these orbitals in terms of the probability of finding an electron in any specified region of space. Often we are interested in the probability of finding the electron within a spherical shell of radius r and thickness dr . This spherical shell can be envisioned as the region between a ball of radius $r + dr$ and another ball of radius r nested inside the first ball. The probability of finding the electron in this spherical shell region is $P(r)dr = r^2 R_{nl}(r)^2 dr$ for a given orbital, where $P(r)$ is called the *radial distribution function* and can be viewed as a type of probability density. The most probable radius at which the electron will be found is specified by the maximum of $P(r)$ and can be found by simple differentiation. For example, the most probable radius for a $1s$ orbital in hydrogen is at $r = a_0$, the Bohr radius, which is 52.9 pm. This distance is called the Bohr radius because it is related to the radius of the lowest energy electron orbit in Bohr's early model of the hydrogen atom, which is considered to be obsolete but still has some qualitatively correct features.

For biological applications, we are interested in going beyond hydrogenic atoms, which have only a single electron, and treating many electrons quantum mechanically. The electronic wavefunctions and energies for molecular systems can be calculated to obtain structural, energetic, and kinetic information. For example, these methods can be used to identify the lowest energy geometry of a molecule, to characterize the strength of a hydrogen-bonding interaction between two molecules, or to calculate the activation energy for a chemical reaction. For such systems with many electrons, we can no longer solve the Schrödinger equation exactly, and we must invoke approximations and develop numerical techniques. The remainder of this chapter will explore these more complicated situations.

8.3 MANY-ELECTRON ATOMS

In the previous section, we discussed the solution of the Schrödinger equation for hydrogenic atoms, which have only a single nucleus and a single electron. In this section, we will discuss many-electron atoms, which have a single nucleus and many electrons. In subsequent sections, we will move on to molecules, which have many nuclei and many electrons. Studying many-electron atoms before moving on to the more complicated case of molecular systems allows us to introduce some key concepts that will also be applicable to molecules. These concepts include electron spin, the orbital approximation, the Aufbau or building-up principle, and the Pauli principle.

For many-electron atoms, the nucleus is assumed to be fixed, and the electrons are moving around this fixed nucleus. For these types of systems, the Schrödinger equation can be written as

$$-\frac{\hbar^2}{2m_e} \sum_{i=1}^N \nabla_i^2 \Psi_n(\mathbf{r}) + U(\mathbf{r}) \Psi_n(\mathbf{r}) = E_n \Psi_n(\mathbf{r}) \quad (8-3)$$

Here m_e is the mass of an electron, and $\mathbf{r}$ represents the coordinates of all of the N electrons in the atom. Analogous to the simpler equation in one dimension given in Eq. (7-5), this equation can be understood as the sum of kinetic energy and potential energy terms on the left side and a total energy term on the right side. The first term on the left side of Eq. (8-3) corresponds to the kinetic energy of all electrons, where $\sum_{i=1}^N \nabla_i^2 \Psi_n(\mathbf{r})$ denotes the sum of second derivatives of the wavefunction with respect to each electronic coordinate. The second term on the left side of Eq. (8-3) corresponds to the potential energy, which includes the electrostatic Coulomb interactions among the electrons and the nucleus.

The potential energy associated with electrostatic interactions is $U(r) = q_1 q_2 / r$ for two particles with charges q_1 and q_2 separated by distance r . Thus, the potential energy for a system with N electrons is

$$U(\mathbf{r}) = - \sum_{i=1}^N \frac{Ze^2}{r_i} + \sum_{i=1}^N \sum_{j < i}^N \frac{e^2}{r_{ij}} \quad (8-4)$$

Here r_i is the distance between electron i and the nucleus, and r_{ij} is the distance between electrons i and j . The first term is the sum of the interactions of all electrons with the nucleus of charge Ze , and the second term is the sum of the interactions of all pairs of electrons. (The double summation notation in the second term means that all unique pairs of electrons are included.) Note that the sign of the electron–nucleus interaction is negative, indicating an attractive interaction that lowers the overall energy, while the sign of the electron–electron interaction is positive, indicating a repulsive interaction that raises the overall energy. All of these terms are included on the left side of Eq. (8-3), leading to a complicated differential equation that cannot be solved analytically.

This problem can be simplified if we neglect the electron–electron interactions. In this case, the left side of Eq. (8-3) can be separated into a sum of terms that each depend on the coordinates of only one electron. We can then assume that each electron occupies an orbital, which is simply another term for a one-electron wavefunction, and the total wavefunction is just the product of these orbitals. When we are describing atoms, these orbitals resemble the hydrogenic atom orbitals that were discussed in the previous section. Thus, we can just place the electrons in the orbitals denoted $1s$, $2s$, $2p$, and so forth to create the many-electron wavefunction. The energy associated with this wavefunction is the sum of the energies of the occupied orbitals. When the electron–electron interactions are included, we can no longer separate Eq. (8-3) in this way, and this solution is only approximate. Nevertheless, it is a useful approximate model for qualitatively describing atomic wavefunctions. For example, the electronic wavefunction for the He atom, which has two electrons, can be described as the product of two $1s$ orbitals, where two electrons occupy the $1s$ orbital, and the associated energy is twice the energy of the $1s$ orbital.

Before proceeding to atoms with more electrons, we need to introduce the additional concept of electron spin. The spin of an electron is an intrinsic angular momentum often described in terms of the motion of the electron about its own axis. Conveniently, the spin can be described similarly to the orbital angular momentum associated with an electron moving around the nucleus, as discussed for hydrogenic atoms in the previous section. Instead of l and m_l , we use s and m_s for the spin quantum number and the spin magnetic quantum number, respectively. Analogous to the orbital angular momentum, the magnitude of the spin angular momentum is $[s(s+1)]^{1/2}\hbar$, and the z -component of the spin angular momentum is $m_s\hbar$. For an electron, the only value of s that is allowed is $s = 1/2$, and the spin can lie in two different orientations corresponding to $m_s = \pm 1/2$. Often the $m_s = +1/2$ state is denoted *spin up* or α , and the $m_s = -1/2$ state is denoted *spin down* or β .

The total electronic wavefunction not only includes the spatial orbitals such as $1s$, $2s$, and so forth, but also includes a spin component. Moreover, the total wavefunction must satisfy the very important Pauli principle, which states that the electronic wavefunction changes sign when the labels of any two electrons are interchanged. This property is also referred to as the *antisymmetry principle* because a function that changes sign when two coordinates or labels are interchanged is called an antisymmetric function. Consider a wavefunction $\Psi(1, 2)$ for two electrons with labels 1 and 2, which denote both the spatial coordinates $\mathbf{r}_1$ and $\mathbf{r}_2$ and the spins of the two electrons. The Pauli principle implies that

$$\Psi(1, 2) = -\Psi(2, 1) \quad (8-5)$$

A familiar consequence of the Pauli principle is the Pauli exclusion principle, which states that no more than two electrons can occupy any given orbital, and two electrons in a single orbital must have paired spins.

Using the principles discussed so far, we can generate approximate electronic wavefunctions for atoms using the Aufbau principle, where *Aufbau* means *building up* in German. As discussed in the previous section, the energies of the

orbitals for hydrogenic atoms are determined by only the principal quantum number, leading to degeneracies for all orbitals with the same principal quantum number, even if the angular momentum quantum number is different. For example, the $2s$, $2p_x$, $2p_y$, and $2p_z$ orbitals all have the same energy. However, for atoms with more than a single electron, the degeneracy between the different angular momentum quantum numbers is broken because of shielding effects. Specifically, the additional electrons exert a shielding effect so that each electron does not feel the full charge of the nucleus but rather feels a slightly lower nuclear charge, leading to a decrease in the electron–nucleus attractive interaction. Because the shapes of s and p orbitals are different, with an electron in an s orbital more likely to be found closer to the nucleus, the shielding effect is less for electrons in s orbitals than for electrons in p orbitals, and the s orbital is lower in energy than the p orbitals for a given principal quantum number. Similar effects influence the relative energies of the orbitals with higher angular momentum quantum numbers.

As a result of these effects, the energy ordering of the atomic orbitals is $1s$, $2s$, $2p$, $3s$, $3p$, $4s$, $3d$, and so forth. According to the Aufbau principle, the electronic wavefunction for an atom is generated by sequentially filling up these orbitals with two electrons per orbital, starting with the lowest energy orbital. To completely define this process for degenerate orbitals, we also need to invoke Hund's rule, which is comprised of two parts. The first part states that the electrons will occupy different orbitals associated with a given energy level before doubly occupying any of them. The second part states that all electrons in the singly occupied orbitals will have parallel spins. The physical basis for the electrons occupying different orbitals is that the energetically unfavorable electron–electron repulsion and shielding effects are typically smaller when electrons are in different orbitals. The physical basis for the unpaired electrons having the same spins will be discussed subsequently in the context of energetically favorable interactions between electrons with parallel spins. Figure 8-3 illustrates how the Aufbau principle and Hund's rule can be used to determine the electronic configuration for the nitrogen atom.

To understand how these principles can be used to create electronic wavefunctions, we consider the example of the He atom, which has a nucleus of charge $+2e$

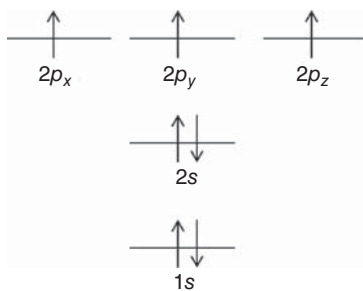


FIGURE 8-3. The electronic configuration for the nitrogen atom, which has seven electrons. The $1s$ and $2s$ orbitals are each doubly occupied, and the three degenerate $2p$ orbitals are singly occupied.

and two electrons. According to the Aufbau principle, the lowest energy configuration corresponds to both electrons occupying the $1s$ orbital given by $\psi_{1s}(\mathbf{r})$. According to the Pauli exclusion principle, one electron must be α (spin up) and the other electron must be β (spin down). Finally, we invoke the antisymmetry principle and create a wavefunction that will change sign when the two electrons are interchanged. The resulting wavefunction is of the form

$$\Psi_{1s^2}(1, 2) = \left(\frac{1}{2}\right)^{1/2} \psi_{1s}(\mathbf{r}_1)\psi_{1s}(\mathbf{r}_2)[\alpha(1)\beta(2) - \alpha(2)\beta(1)] \quad (8-6)$$

Here the labels 1 and 2 refer to the first and second electron, respectively, with spatial coordinates given by $\mathbf{r}_1$ and $\mathbf{r}_2$, respectively. Note that if you interchange the labels 1 and 2 in Eq. (8-6), the wavefunction changes sign, thereby satisfying Eq. (8-5). This wavefunction is often denoted $1s^2$ because there are two electrons in the $1s$ orbital.

For two-electron systems, we are able to construct electronic wavefunctions that satisfy the Pauli principle using mathematical intuition. Unfortunately, this approach does not work for systems with more electrons. In these cases, we need to use more advanced mathematical methods to construct the electronic wavefunctions. In addition, up to this point we have assumed that the orbitals are simply the hydrogenic atom orbitals, but usually these are not the optimal orbitals. In the next two sections, we will discuss numerical techniques that can be used to construct electronic wavefunctions that satisfy the Pauli principle and generate optimal orbitals that lead to lower energies. The methods described in the subsequent sections will be useful for determining electronic wavefunctions for many-electron atoms as well as molecules.

8.4 BORN–OPPENHEIMER APPROXIMATION

For atoms, we have assumed that the nucleus is fixed and the electrons move around this fixed nucleus. This assumption is reasonable in that we can define the coordinate system for the electrons such that the much heavier nucleus is at the origin. The situation becomes more complicated when there are many nuclei, such as in molecules. Solving the Schrödinger equation for all electrons and all nuclei in a large macromolecule is not computationally practical, even using advanced numerical techniques. A key simplification that is central to most of quantum chemistry is the Born–Oppenheimer approximation. This approximation is based on the fact that nuclei are much heavier than electrons and therefore move more slowly. As a result, the electrons can be considered to be moving in the field of fixed nuclei. In other words, the electrons are assumed to respond instantaneously to the motion of the nuclei. A helpful analogy is the motion of a horse and the swarm of flies around her head: as the horse walks along, the flies move much faster and appear to respond instantaneously to the horse's motion and thus to move as a cloud surrounding her head. Similarly, as the nuclei in a molecule move, the electrons respond so quickly that we can treat the electrons as moving in the field of fixed nuclei at each nuclear geometry.

We can also describe the Born–Oppenheimer approximation in the framework of the Schrödinger equation. For a system made up of nuclei and electrons, this equation includes kinetic energy terms for the electrons and the nuclei, as well as potential energy terms for the interactions among the electrons and nuclei. According to the Born–Oppenheimer approximation, the nuclei are moving so slowly that the kinetic energy of the nuclei and the nuclear–nuclear repulsion term are considered separately. In this case, we only need to solve the electronic Schrödinger equation, which includes the kinetic energy of the electrons and the electron–electron and electron–nucleus potential energy interactions, for fixed nuclear positions. This electronic problem is much easier to solve and can be written as

$$-\frac{\hbar^2}{2m_e} \sum_{i=1}^N \nabla_i^2 \Psi_n(\mathbf{r}; \mathbf{R}) + U(\mathbf{r}, \mathbf{R}) \Psi_n(\mathbf{r}; \mathbf{R}) = E_n(\mathbf{R}) \Psi_n(\mathbf{r}; \mathbf{R}) \quad (8-7)$$

Here all of the electronic coordinates are denoted by $\mathbf{r}$, and all of the nuclear coordinates are denoted by $\mathbf{R}$. The potential energy term for a system of M nuclei and N electrons is of the form

$$U(\mathbf{r}, \mathbf{R}) = - \sum_{I=1}^M \sum_{i=1}^N \frac{Z_I e^2}{r_{iI}} + \sum_{i=1}^N \sum_{j < i}^N \frac{e^2}{r_{ij}} \quad (8-8)$$

This expression is very similar to the one used for many-electron atoms (Eq. (8-4)) except that the first term is summed over all nuclei, where r_{iI} is the distance between electron i and nucleus I . Equation (8-7) is very similar to the Schrödinger equation for many-electron atoms (Eq. (8-3)) except for the dependence on the nuclear coordinates $\mathbf{R}$. The semicolon in the wavefunction indicates that it depends explicitly on $\mathbf{r}$ but only *parametrically* on $\mathbf{R}$. This means that for each set of fixed nuclear coordinates $\mathbf{R}$, we solve this equation to obtain the electronic wavefunctions and energies for that particular $\mathbf{R}$, that is, $\Psi_n(\mathbf{r}; \mathbf{R})$ and $E_n(\mathbf{R})$. We repeat this procedure for all possible sets of nuclear coordinates $\mathbf{R}$ to obtain the wavefunctions and energies that depend parametrically on $\mathbf{R}$.

According to the Born–Oppenheimer approximation, the potential energy function for the nuclei is given by

$$U(\mathbf{R}) = E_0(\mathbf{R}) + \sum_{I=1}^M \sum_{J < I}^M \frac{Z_I Z_J e^2}{R_{IJ}} \quad (8-9)$$

The first term on the right side of this equation is the energy of the lowest energy state ($n = 0$) obtained by solving the electronic Schrödinger equation given in Eq. (8-7) for all nuclear coordinates $\mathbf{R}$, and the second term on the right side corresponds to the repulsive interactions between all pairs of nuclei. The double summation notation in this second term indicates that we are summing over all unique pairs of nuclei, where R_{IJ} is the distance between nucleus I and nucleus J . The simplest example is a diatomic molecule, where the only nuclear coordinate that corresponds to a vibration

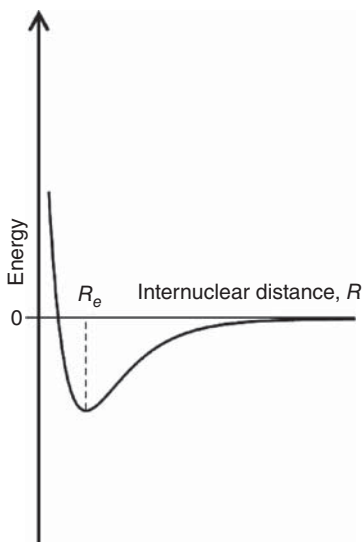


FIGURE 8-4. The potential energy curve, $U(R)$, for a diatomic molecule. The coordinate R corresponds to the internuclear distance, and the minimum of this curve corresponds to the equilibrium bond length R_e .

is the internuclear distance. (The other five degrees of freedom are associated with translations and rotations of the diatomic molecule.) In this case, we can solve Eq. (8-7) for each internuclear distance R to obtain $E_0(R)$, and the second term in Eq. (8-9) is comprised of the repulsion interaction between the two nuclei. Figure 8-4 depicts the potential energy curve, $U(R)$, that is generated by plotting the potential energy given in Eq. (8-9) as a function of the internuclear distance R . For molecules with more degrees of freedom, the potential energy surface will depend on all of the nuclear vibrational coordinates, after removing the translations and rotations.

8.5 MOLECULAR ORBITAL THEORY

In Section 8.3, we were able to construct the electronic wavefunction for the He atom using mathematical intuition. If we want to study molecules, however, we need to use more advanced techniques. One of the most powerful techniques for this purpose is molecular orbital (MO) theory. In MO theory, we construct MOs from a linear combination of atomic orbitals, which are one-electron wavefunctions typically centered on the nuclei. The underlying concepts of the Aufbau principle and the Pauli principle still hold, but in this case, the electrons occupy MOs rather than atomic orbitals. Thus, we fill up the MOs with pairs of electrons starting with the lowest energy MO.

The easiest way to understand MO theory is to consider a homonuclear diatomic molecule, starting with H_2 . Assume that we have two $1s$ atomic orbitals, with one centered on hydrogen nucleus A and the other centered on hydrogen nucleus B.

We denote these two atomic orbitals $1s_A$ and $1s_B$, respectively. We can then create bonding and antibonding MOs from these atomic orbitals by taking the sum and the difference, respectively:

$$\begin{aligned}\psi_{\text{bond}} &= N_+(1s_A + 1s_B) \\ \psi_{\text{anti}} &= N_-(1s_A - 1s_B)\end{aligned}\quad (8-10)$$

where the dependence on the electronic coordinate is omitted for notational simplicity. The constant prefactors, N_+ and N_- , ensure that the total probability of the electron being anywhere in space is unity for each of these MOs. Recall from Chapter 7 that the probability of finding the electron at any point in space is proportional to the square of the wavefunction, which is called the probability density. The probability densities corresponding to the wavefunctions given in Eq. (8-10) are:

$$\begin{aligned}\psi_{\text{bond}}^2 &= N_+^2[1s_A^2 + 1s_B^2 + 2(1s_A 1s_B)] \\ \psi_{\text{anti}}^2 &= N_-^2[1s_A^2 + 1s_B^2 - 2(1s_A 1s_B)]\end{aligned}\quad (8-11)$$

The terms $1s_A^2$ and $1s_B^2$ are the probability densities associated with the atomic orbitals on hydrogen nucleus A and B, respectively. The additional term, $\pm 2(1s_A 1s_B)$, is denoted the *overlap density* and is positive for the bonding MO and negative for the antibonding MO, corresponding to constructive and destructive interference, respectively. The concepts of constructive and destructive interference were discussed in the context of light waves in Chapter 7. In the present case, constructive interference refers to the two atomic orbitals having the same sign in the internuclear region, leading to positive overlap density, and destructive interference refers to the two atomic orbitals having opposite signs in the internuclear region, leading to negative overlap density.

The overlap density term is important because it reflects the probability of finding the electron in the internuclear region between the two hydrogen nuclei A and B. As shown in Figure 8-5, when this term is positive, as in the bonding MO, there is extra probability in this internuclear region. In this case, the total probability of finding the electron between the two nuclei is greater than it would be if the electron were confined to one of the atomic orbitals. When this term is negative, as in the antibonding MO, there is a diminished probability in this internuclear region. In this case, the total probability of finding the electron between the two nuclei is smaller than it would be if the electron were confined to one of the atomic orbitals. This is a general characteristic of bonding and antibonding orbitals: bonds form when electronic density accumulates between the nuclei due to constructive interference of atomic orbitals.

Another important property of bonding and antibonding orbitals such as those given in Eq. (8-10) is that the energy of the bonding orbital is lower than the energy of the antibonding orbital. Moreover, the energy of the bonding orbital is lower than the energies of the individual atomic orbitals, whereas the energy of the antibonding orbital is greater than the energies of the individual atomic orbitals. Furthermore,

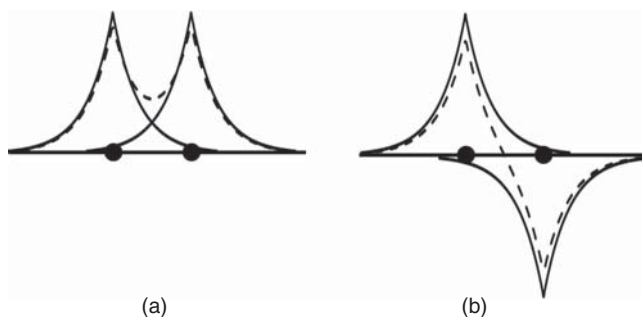


FIGURE 8-5. Schematic depiction of the bonding and antibonding MOs for a homonuclear diatomic molecule, where the two nuclei are represented by filled circles. The solid lines depict the $1s$ atomic orbitals, and the dashed lines depict the MOs defined in Eq. (8-10). Figure 8-5(a) depicts constructive interference between two $1s$ atomic orbitals, associated with a bonding MO. Figure 8-5(b) depicts destructive interference between two $1s$ atomic orbitals, associated with an antibonding MO. Note that the bonding orbital has increased electronic density in the internuclear region, whereas the antibonding orbital has decreased electronic density in this region and becomes zero at the midpoint.

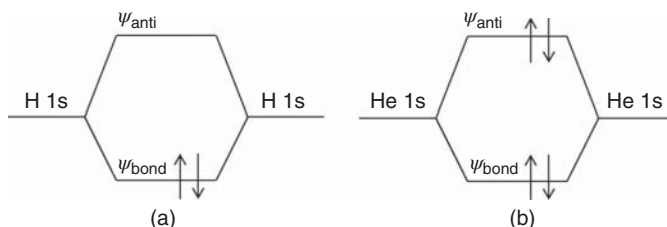


FIGURE 8-6. MO diagrams for H_2 (a) and He_2 (b) diatomic molecules. Note that antibonding is slightly more energetically unfavorable than bonding is energetically favorable. H_2 is a stable species because two electrons are in the bonding MO, which is lower in energy than the separated atoms. He_2 is not a stable species because the two additional electrons are in the antibonding MO, leading to a higher energy than the separated atoms.

antibonding is slightly more energetically unfavorable than bonding is energetically favorable, as depicted in Figure 8-6 for the H_2 molecule. These properties can be shown mathematically by substituting the wavefunctions in Eq. (8-10) into the Schrödinger equation (Eq. (8-7)) to determine the associated energies. The mathematics is somewhat involved, however, so we will not pursue this matter further. We can use the same Aufbau principle for H_2 that we used for atomic wavefunctions and place the two electrons with paired spins in the lowest energy MO, which in this case is the bonding MO. As depicted in Figure 8-6, this configuration is lower in energy than the separated H atoms, so H_2 is a stable species. On the other hand, if we consider He_2 , which has 4 electrons, the extra two electrons would be placed

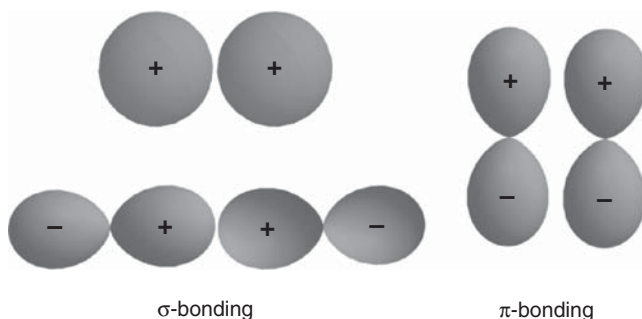


FIGURE 8-7. Schematic depiction of σ MOs comprised of either s or p orbitals with cylindrical symmetry about the interatomic axis and a π MO comprised of p orbitals that are perpendicular to the interatomic axis for a diatomic molecule. The antibonding MOs correspond to changing the signs for one of the orbitals in each pair.

in the antibonding orbital, and the overall energy would be greater than that of the separated He atoms (Figure 8-6). As a result, He_2 is not a stable species and is not found in nature. However, if we remove one electron, creating the He_2^+ cation, there are two electrons in the bonding MO and one electron in the antibonding MO, and the overall energy is lower than that of the separated He atoms. Thus, the He_2^+ molecule is a stable species.

The same process can be used to describe homonuclear diatomic molecules with more electrons. In general, we can build N MOs from N atomic orbitals. For this purpose, it is convenient to introduce σ and π MOs: σ orbitals have cylindrical symmetry about the internuclear axis, and π orbitals are perpendicular to the internuclear axis. A schematic depiction of σ and π MOs is given in Figure 8-7. In MO theory, all orbitals of the appropriate symmetry will contribute to those types of MOs, but often we can use a simpler description if orbitals are well-separated in energy. An illustrative example is O_2 , where the two oxygen atoms are placed along the z -axis. We consider only the valence orbitals, defined to be the $2s$ and $2p$ orbitals, because the $1s$ core orbitals are not expected to participate in bonding. In terms of the σ MOs, the $2s$ orbitals on each oxygen atom can form bonding and antibonding orbitals, analogous to those given in Eq. (8-10), and the $2p_z$ orbitals on each oxygen atom can form bonding and antibonding orbitals, leading to a total of two bonding and two antibonding σ MOs. In terms of the π MOs, the $2p_x$ and $2p_y$ orbitals on the two oxygen atoms can combine to create two bonding and two antibonding π MOs. The resulting MO diagram for O_2 is shown in Fig. 8-8, where the subscripts g and u denote the German words *gerade* and *ungerade*, translating to even and odd, respectively, for the symmetry of the orbitals.

In general, the MO diagrams for diatomic molecules can be constructed in this manner. Depending on the relative energies of the $2s$ and $2p$ atomic orbitals, the relative energies of the MOs can shift. In particular, the order of the $2\sigma_g$ and $1\pi_u$ orbitals changes for certain diatomic molecules, such as N_2 . Moreover, for heteronuclear diatomic molecules, such as HF, the relative weightings of the atomic orbitals on the

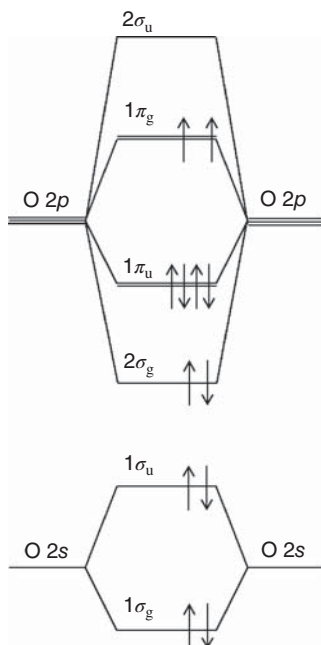


FIGURE 8-8. Schematic MO diagram for the O_2 diatomic molecule. An analogous diagram can be used for F_2 , but the order of the $2\sigma_g$ and $1\pi_u$ orbitals should be switched for Li_2 , Be_2 , B_2 , C_2 , and N_2 .

two atoms will be different, rather than equally combined as in Eq. (8-10), and the bonds become polar. For HF, the bonding orbitals have greater contributions from atomic orbitals on F, and the antibonding orbitals have greater contributions from atomic orbitals on H.

8.6 HARTREE-FOCK THEORY AND BEYOND

The creation of bonding and antibonding MOs for diatomic molecules is relatively straightforward. When we get to more complicated molecules, however, we can no longer rely on simple combinations of atomic orbitals. Instead, we need to have a more general approach for generating the electronic wavefunctions. An essential requirement is that the wavefunctions must be antisymmetric with respect to interchange of any two electrons, as dictated by the Pauli principle. For this purpose, we can use a Slater determinant form of the wavefunction. The mathematical details of Slater determinants are beyond the scope of this book, and we will simply view this as a way to combine occupied MOs with appropriate spin components to generate an antisymmetric wavefunction. This is the basis of Hartree-Fock theory (2).

For large molecules, we cannot simply guess the atomic orbitals that will be combined to make each MO. Instead, we define basis functions that are combined to

construct all of the MOs. In practice, we choose a basis set comprised of N_{bf} basis functions, denoted $\varphi_{\mu}(\mathbf{r}_1)$. The basis functions are often chosen to have the form of the s , p , and d hydrogenic atom orbitals and are centered on all of the nuclei in the system. The size and flexibility of the basis set needs to be sufficient to adequately describe the total electronic wavefunction of the system. Each MO, denoted $\psi_i(\mathbf{r}_1)$, is described as a linear combination of these N_{bf} basis functions:

$$\psi_i(\mathbf{r}_1) = \sum_{\mu=1}^{N_{\text{bf}}} c_{i\mu} \varphi_{\mu}(\mathbf{r}_1) \quad (8-12)$$

Here the MO coefficients, $c_{i\mu}$, tell us how the basis functions are combined for the MO $\psi_i(\mathbf{r}_1)$. The MOs given in Eq. (8-12) can be combined with spin components to obtain the total antisymmetric wavefunction for a particular molecule.

We can calculate the total energy by substituting the wavefunction for a specified set of MOs into the Schrödinger equation given in Eq. (8-7). Consider a system of N electrons with $N/2$ MOs that are each occupied with two electrons of opposite spin. For this system, the Hartree–Fock energy is:

$$E = 2 \sum_{i=1}^{N/2} h_i + \sum_{i=1}^{N/2} \sum_{j=1}^{N/2} [2J_{ij} - K_{ij}] \quad (8-13)$$

Here the summations are over all doubly occupied MOs. The first term is a sum over h_i , which includes the kinetic energy and the electron–nucleus interaction for the i th MO. The electron–nucleus interaction can be viewed as the classical Coulombic attraction between the electronic charge cloud with probability density $|\psi_i(\mathbf{r}_1)|^2$ and all of the nuclei. The second term arises from the electron–electron interactions and therefore includes a double summation representing all pairs of electrons. This electron–electron interaction term is comprised of two contributions: the Coulomb terms, J_{ij} , and the exchange terms, K_{ij} . The Coulomb term is associated with the classical Coulomb repulsion between electronic charge clouds with probability densities $|\psi_i(\mathbf{r}_1)|^2$ and $|\psi_j(\mathbf{r}_2)|^2$.

The exchange term does not have a simple classical interpretation and arises from interactions between electrons with the same spin (parallel spins). Because the exchange term is positive and appears with a negative sign in the total energy given in Eq. (8-13), the exchange interactions lower the energy of the system. Thus, the exchange term can be viewed as the energetically favorable contribution from correlated motion between electrons with parallel spins. Moreover, analysis of the exchange terms for systems with unpaired spins can be used to justify the second part of Hund’s rule stating that the lowest energy wavefunction corresponds to a configuration in which all electrons in singly occupied orbitals have parallel spins.

The next challenge is to determine the coefficients in Eq. (8-12) for a particular molecule. For this purpose, we invoke an important theorem of quantum mechanics: the variation principle. The variation principle states that if an arbitrary wavefunction is used to calculate the energy, the calculated value is always greater than or equal to

the true energy. Given this theorem, the goal is to find the MO coefficients that will lead to the lowest energy of the system. The general procedure, which is called the variational method, is to substitute the wavefunction into the Schrödinger equation (Eq. (8-7)) and solve for the coefficients that minimize the energy. When all electrons are paired, the energy is given by Eq. (8-13). The variational method produces a set of coupled differential equations, which are called the Hartree–Fock equations. These equations can be used to solve for the coefficients that minimize the energy. This procedure is called a self-consistent-field (SCF) method because the equations need to be solved iteratively: an initial guess for the MO coefficients leads to a new set of coefficients, which are then used to generate another new set of coefficients, and so forth until the MO coefficients no longer change significantly and are considered to be converged. We will not go into the details of this procedure, but a wide range of numerical methods have been developed for this purpose.

The variational method can be illustrated by considering the H_2 molecule discussed earlier. Let us choose a basis set comprised of the two basis functions, $1s_A$ and $1s_B$, defined to be $1s$ atomic orbitals centered on each hydrogen nucleus. We assume that the two electrons are paired in the same MO, which we denote $\psi_1(\mathbf{r}_1)$. Then we expand this MO in terms of the two basis functions, $1s_A$ and $1s_B$, to obtain

$$\psi_1 = c_{1A}1s_A + c_{1B}1s_B \quad (8-14)$$

where the dependence on the electronic coordinate is omitted for notational simplicity. According to the variational principle, our goal is to determine the coefficients, c_{1A} and c_{1B} , of the two basis functions that will lead to the lowest energy of the system. Solving the Hartree–Fock equations resulting from the variational method, we find that the two coefficients are equal, $c_{1A} = c_{1B}$, leading to the MO given by ψ_{bond} in Eq. (8-10). For this simple problem, we were able to determine the lowest energy wavefunction from symmetry arguments. For systems with more electrons and more complicated basis sets, however, we need to use the numerical framework provided by Hartree–Fock theory.

Relying on physical intuition, we might choose to use the hydrogenic atom orbitals as the basis functions in Eq. (8-12). However, the form of these atomic orbitals is not conducive to the efficient calculation of integrals required for the variational method described earlier. Instead, typically the basis functions are chosen to be linear combinations of simpler functions designed to mimic the hydrogenic atom orbitals. According to the variational method, as we add more basis functions, the energy of the system will decrease and get closer to the true energy. The best Hartree–Fock energy is determined by increasing the basis set until the energy no longer changes significantly, or is converged to the desired level of accuracy. John Pople won the Nobel Prize in 1998 for his development of these types of computational methods in quantum chemistry.

So far, we have discussed *ab initio* methods, which are rigorously defined to be methods based on first principles of quantum mechanics and do not involve any experimental input. For applications to large biomolecules, *ab initio* methods such as Hartree–Fock theory are computationally expensive, especially in conjunction with

large basis sets. The main expense is the calculation of the integrals required for determination of the Coulomb and exchange terms in the energy expression given by Eq. (8-13). For this reason, a variety of semiempirical methods have been developed for the study of larger systems. In semiempirical methods, the integrals are not calculated exactly but rather are estimated using a prescribed set of physically motivated rules. Many of the integrals are set to zero and therefore do not need to be calculated at all. Other integrals are approximated by simple functional forms with parameters that are fit to reproduce experimental data for a set of molecules. Semiempirical methods are much more computationally efficient than *ab initio* methods. The disadvantage of semiempirical methods, however, is that they tend to be reliable only for molecules that are closely related to the molecules used in the parameterization procedure. Thus, semiempirical methods must be used with care.

A variety of *ab initio* methods that have been developed recently provide a more rigorous but computationally expensive strategy. The difference between the exact energy and the Hartree–Fock energy is called the electron correlation energy. Advanced *ab initio* methods have been designed to account for some of the electron correlation energy and are available in a wide range of popular quantum chemistry programs. However, such *ab initio* methods are not directly applicable to large biomolecules due to the computational expense. An alternative, more efficient approach is density functional theory (DFT), which is discussed in the next section.

8.7 DENSITY FUNCTIONAL THEORY

DFT is a very powerful method in quantum chemistry for studying large molecules because it is able to strike a balance between accuracy and computational efficiency (3). In DFT, the Schrödinger equation is represented in terms of the electronic density, $\rho(\mathbf{r})$, where $\mathbf{r}$ is a three-dimensional spatial coordinate, instead of the $3N$ -dimensional total electronic wavefunction. The electronic density is related to the probability of finding the electrons anywhere in space. In particular, $\rho(\mathbf{r})dV$ tells us the probability of finding electrons in a volume element dV centered at a point $\mathbf{r}$ in space. If we integrate the electronic density over all space, we obtain the number of electrons in the system. The underlying principle of DFT is that the total energy of the system is a functional of the electronic density. A functional is defined to be a function of another function: in this case, the energy is a function of the electronic density, which is a function of position. Thus, the energy functional is expressed as $E[\rho(\mathbf{r})]$. If the electronic density in three-dimensional space is provided, this functional provides the energy of the system associated with that density.

In principle, DFT is an exact theory. In practice, however, we need to use approximate methods because we do not know the form of the exact energy functional. Most current implementations of DFT for chemical and biological systems rely on the Kohn–Sham formalism, which closely resembles Hartree–Fock theory in spirit, except the MOs are called Kohn–Sham orbitals. The electronic density, $\rho(\mathbf{r})$, can be obtained by summing over the probability densities of the N occupied Kohn–Sham orbitals. The kinetic energy terms, as well as the potential energy terms associated

with the classical Coulomb interactions among the electrons and nuclei, are included in a manner analogous to Hartree–Fock theory. In DFT, however, additional contributions to the energy are included through the use of a *density functional*.

A detailed discussion of the origin and form of density functionals is beyond the scope of this book. To understand the basic purpose of density functionals, we point out that they account for both electron exchange and electron correlation effects. Recall that the Hartree–Fock energy includes contributions from exchange terms, which are associated with the correlated motion between electrons with parallel spins. In the energy expression given by Eq. (8-13), the classical electron–electron Coulomb terms are denoted by J_{ij} , and the exchange terms are denoted by K_{ij} . Many density functionals replace all or a portion of the Hartree–Fock exchange terms with approximate terms. In addition, density functionals add electron correlation, which is defined to be the difference between the exact energy and the Hartree–Fock energy and thus is completely missing in Hartree–Fock theory. A large number of different density functionals have been developed over the years. The most widely used density functionals contain parameters that have been determined by fitting to known properties, such as geometries, energies, frequencies, and thermochemical properties of large sets of molecules. The resulting functionals tend to be reliable only for molecules that are related to those in the set used in the parameterization procedure. The development of more general functionals based on rigorous derivations is a direction of continued research.

In practice, we need to consider several important points when performing DFT calculations on macromolecules. First, we need to choose a sufficiently large and flexible basis set to adequately describe the system. Second, we need to choose a density functional that is appropriate for this type of system. For example, certain density functionals are better at describing hydrogen-bonding interactions, while other density functionals are better at describing transition metals. In many cases, performing calculations with multiple functionals is useful for comparative purposes. Keeping all of this in mind, DFT has provided significant insights into chemical, biological, and physical systems. Along with John Pople, Walter Kohn won the Nobel prize in 1998 for the development of DFT.

8.8 QUANTUM CHEMISTRY OF BIOLOGICAL SYSTEMS

In this section, we explain how to use the quantum chemistry methods discussed earlier to better understand biological systems. According to the Born–Oppenheimer approximation, the nuclei move on a potential energy surface defined by $U(\mathbf{R})$ given in Eq. (8-9). This potential energy surface is generated by solving the electronic Schrödinger equation (Eq. (8-7)) to obtain the electronic energy for all nuclear configurations and adding the nuclear–nuclear interaction term. In Chapter 7, we discussed the concept of normal modes of vibration that characterize the motion of a molecule. In general, each normal mode is a combination of the motions of many different bonds, although certain normal modes are dominated by the motion of one or a few chemical bonds. In quantum chemistry calculations, it is often convenient to

characterize points on the potential energy surface in terms of these normal mode coordinates.

In general, the potential energy surface is a complex multidimensional surface. For a nonlinear molecule comprised of N nuclei, there are $3N-6$ vibrational degrees of freedom after removing the three translational and three rotational degrees of freedom for the overall molecule, leading to a $3N-6$ -dimensional potential energy surface. Note that a diatomic molecule is much simpler because the potential energy surface is only one-dimensional, as depicted in Figure 8-4. In this simpler case, it is easy to generate the entire one-dimensional curve as a function of the internuclear bond length R . From this curve, we are able to determine the equilibrium bond length, which is the value of R at the minimum of the curve, and the frequency of the vibrational mode, which is related to the second derivative at the minimum of the curve. The situation becomes much more complicated for larger molecules, which can have hundreds of degrees of freedom, and we can no longer visualize the potential energy surface in this way.

Clearly, generating the entire multidimensional surface for a large molecule is not practical. Instead, we focus on identifying and characterizing the important points on this surface. Figure 8-9 depicts a two-dimensional potential energy surface to illustrate the types of points that are of interest. To determine the stable conformations of a molecule, we want to find the minima on the potential energy surface because the forces on all nuclei vanish at the minima. For large molecules, typically there are many different minima on the potential energy surface. Although the most stable

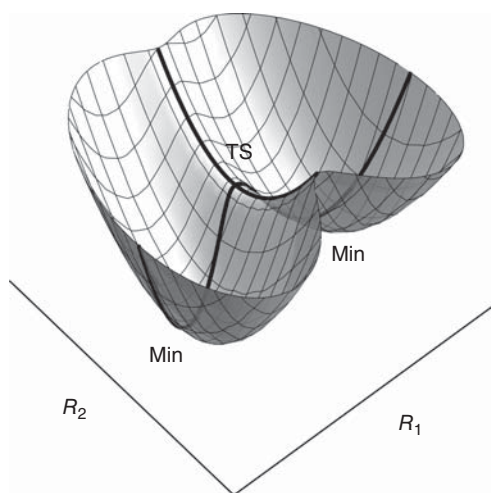


FIGURE 8-9. Two-dimensional potential energy surface illustrating two minima, denoted as Min, and a transition state, denoted as TS. The coordinates R_1 and R_2 are chosen to be the normal modes at the TS, which is a first-order saddle point, as shown by the negative curvature along R_1 and the positive curvature along R_2 (thick dark lines). The minimum energy path from the TS down to the two minima is shown by the thick dark line associated with negative curvature at the TS.

conformation corresponds to the *global minimum* (the minimum of lowest energy), the other minima, called *local minima*, may also represent stable conformations that are chemically relevant. Often we are interested in the relative energies of these stable conformations because they provide information about the relative populations of these species.

In practice, the minima on multidimensional potential energy surfaces can be found using computer algorithms designed to minimize the energy as a function of the nuclear coordinates. Mathematically, a minimum is a specific type of stationary point, which is defined as a point for which the derivative of the energy with respect to all nuclear coordinates is zero. However, not all stationary points are minima. To determine which stationary points are minima, we need to calculate the second derivatives of the energy with respect to the nuclear coordinates. The characterization of stationary points is more straightforward in terms of the normal mode coordinates. When the second derivatives are positive along all coordinates, the stationary point is a minimum. When the second derivatives are positive along all normal mode coordinates except for a single normal mode coordinate, which has a negative second derivative, the stationary point is called a first-order saddle point because in two dimensions, it resembles a saddle that curves up in one direction and curves down in the other direction (Figure 8-9). This type of saddle point is also called a transition state (TS), which provides information about the barrier for a chemical reaction. Often a chemical reaction is assumed to follow the *minimum energy path*, which is defined to be the path of steepest descent from the TS down to the two minima. Figure 8-9 depicts the minimum energy path from a TS down to two minima for a two-dimensional potential energy surface as a function of the normal mode coordinates at the TS. However, the character of the normal mode coordinates may change during a chemical reaction, as illustrated in Figure 8-9. The second derivatives of the energy with respect to the normal mode coordinates are also associated with the frequencies of the vibrational normal modes at the stationary points.

So far, we have talked about minimizing the energy to identify stable species. From the earlier chapters on thermodynamics, however, we learned that stable species actually correspond to the minimum Gibbs energy, which includes entropic effects. (Note that the Gibbs energies are implicitly assumed to be standard Gibbs energies in this context.) Because entropic contributions are often similar for different conformations of a particular molecule, the errors associated with minimizing the energy rather than the Gibbs energy are usually negligible. However, the energies at stationary points can be corrected to include entropic effects using statistical mechanics methods that will be discussed in Chapter 15. The entropic effects due to the molecular vibrations are calculated from the frequencies of the normal modes using the harmonic oscillator model, and the entropic effects due to the translations and rotations are calculated using other models. In addition, the energies can be corrected for zero-point energy effects that were discussed in Chapter 7. The zero-point energy contributions are typically calculated from the frequencies of the vibrational normal modes using the harmonic oscillator model. These types of approaches can be used to calculate thermodynamic quantities with quantum chemistry methods.

In addition to calculating thermodynamic properties, quantum chemistry can also be used to investigate the kinetics of chemical reactions. As mentioned earlier, if a TS is found, then the path of steepest descent can be followed from that TS down to the minima associated with the reactant and product of the chemical reaction. Subsequently, we can add entropic and zero-point energy contributions to the energies of the stationary points and calculate the Gibbs energy of activation, which is the difference in Gibbs energies of the TS and reactant, and the Gibbs energy of reaction, which is the difference in Gibbs energies of the product and reactant. The Gibbs energy of activation can be used to calculate the rate constant with a method such as TS theory, which was discussed in Chapter 5. The Gibbs energy of reaction can be used to calculate the equilibrium constant for the chemical reaction and thereby to determine the relative populations of the reactant and product species at equilibrium.

Most biological reactions occur in solution, and solvent effects can be very important. The effects of solvent can be included in quantum chemistry calculations in several different ways. A popular approach is to treat the solvent as a dielectric continuum surrounding the molecule, including the electrostatic interaction between the molecule and the solvent in the total energy. A disadvantage of this approach is that it cannot describe specific hydrogen-bonding interactions between solvent and solute molecules. Another option is to include a small number of solvent molecules explicitly in the calculation, forming a cluster comprised of the solute and several solvent molecules. A disadvantage of this approach is that the results can depend on where these solvent molecules are placed in the initial configuration because the system can become stuck in local minima. Other approaches that rely on mixing quantum mechanical methods with classical mechanical methods will be discussed in Chapter 16.

Many biological systems are too large to be studied with quantum chemistry methods. In these cases, the full system can be modeled by a smaller portion of the system. For example, an enzyme reaction may be studied by performing a quantum chemistry calculation on only the active site. Unfortunately, if the geometry of the active site of an enzyme is optimized in isolation, the resulting structure could be nonphysical because the remainder of the enzyme restricts the motion of the active site in a way that is not easily replicated in the absence of the enzyme. One approach is to constrain certain atoms in the active site model to positions that are known from experimental structural studies and optimize only the positions of the remaining atoms. Unfortunately, the results from this approach depend on the choice of which atoms to constrain and therefore are somewhat arbitrary. A more reliable and objective approach is to use a hybrid method in which only the active site is treated quantum mechanically, but the remainder of the enzyme is still included in a more approximate way. These types of hybrid approaches will be discussed in Chapter 16.

To illustrate the wide range of information provided by quantum chemistry calculations, we discuss a few select studies of nucleic acid bases and base pairs using *ab initio* methods or DFT. We emphasize that the absolute energies obtained with these types of quantum chemistry methods are not meaningful. Instead, we can only compare energies for systems comprised of the same atoms calculated at the same level of theory. For example, such comparisons require the use of the same basis set

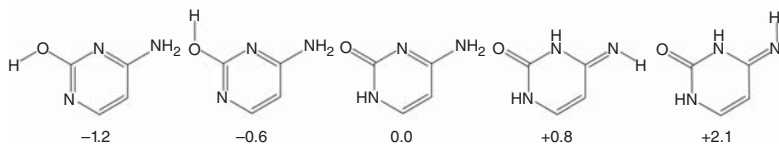


FIGURE 8-10. Low-energy tautomers or isomers of cytosine and the relative gas phase energies (in kcal/mol) with respect to the canonical keto-amino form calculated with an *ab initio* method. Negative energies indicate tautomers more stable than the canonical form, and positive energies indicate tautomers less stable than the canonical form in the gas phase. Data from Ref. 4.

and the same density functional within DFT. Thus, chemical and biological applications of quantum chemistry usually focus on calculating properties such as relative energies of isomers, interaction energies relative to separated species, or relative thermochemical properties of a system undergoing a chemical reaction. Structural information, including bond lengths and angles, can also be obtained from these types of calculations.

For nucleic acid bases, geometry optimizations with high-level *ab initio* methods have been used to determine the structures, dipole moments, and proton affinities of cytosine, adenine, guanine, thymine, and uracil. These calculations have also provided information about the relative energies of tautomers due to internal proton transfers within such nucleic acid bases. For example, the tautomers of cytosine were determined to have the gas phase energies shown in Figure 8-10. These calculated energies indicate that all of these tautomers are accessible at room temperature and therefore agree with experiments showing that all species can coexist in the gas phase. Solvation was found to strongly influence the relative energies of these tautomers, stabilizing the canonical keto-amino form over the imino form by ~ 7 kcal/mol, also in agreement with experimental estimates. As a result, the imino tautomer is not observed often in biomolecules.

The interaction energies of DNA base pairs have also been calculated with quantum chemistry methods. These interaction energies are defined to be the difference between the energy of the base pair and the energies of the isolated bases. The accurate description of these types of interaction energies requires high-level *ab initio* methods that go beyond Hartree–Fock theory. These types of *ab initio* methods indicate that the interaction energy is almost twice as large for the guanine–cytosine (GC) Watson–Crick base pair than for the adenine–thymine (AT) Watson–Crick base pair, as depicted in Figure 8-11. The calculations suggest two reasons for the stronger interaction energy of the GC base pair. The first reason is the different number of hydrogen bonds in the two base pairs, with three hydrogen bonds in the GC pair and only two hydrogen bonds in the AC pair. The second reason is the stronger polarity of guanine and cytosine compared to thymine and adenine, as indicated by the calculated dipole moments of the individual bases. The aromatic base stacking interactions, as well as the interactions of metal cations with bases and base pairs, have also been investigated with quantum chemistry methods. Overall, these types of calculations have provided insights about fundamental interactions in nucleic acids.

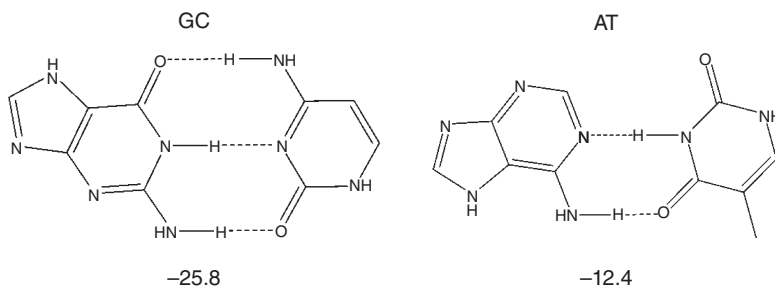


FIGURE 8-11. Hydrogen-bonded GC and AT Watson-Crick base pairs and interaction energies (kcal/mol) calculated with an *ab initio* method. Negative interaction energies indicate that the base pair is more stable than the separated bases, and the lower energy for GC than for AT indicates a stronger interaction energy for GC. Data from Ref. 4.

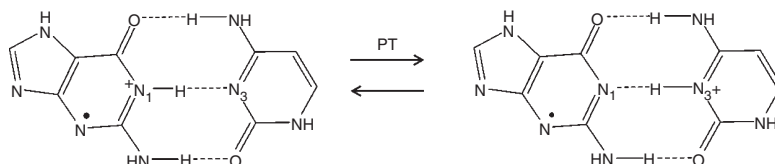


FIGURE 8-12. Schematic depiction of the proton transfer reaction in the GC radical cation base pair.

To illustrate a chemical reaction, we focus on proton transfer in the GC radical cation base pair. Upon one-electron oxidation of guanine in a GC base pair, experiments suggest that a proton is transferred from N1 of the guanine radical cation to N3 of the cytosine. This reaction, which is depicted in Figure 8-12, was studied with DFT in the gas phase and in the presence of eleven explicit water molecules. In both cases, the reactant and product minima corresponding to the proton on the N1 of G or the N3 of C, respectively, and the TS corresponding to the proton midway between the N1 of G and the N3 of C were found. The zero-point energy and entropic contributions were calculated to generate the Gibbs energies for this reaction. A one-dimensional representation of the Gibbs energy surface for the proton transfer reaction in the solvated system is depicted in Figure 8-13. In the gas phase, the proton transfer reaction is not thermodynamically favored, and the Gibbs energy of reaction is 1.38 kcal/mol. In the presence of 11 water molecules, however, the proton transfer is thermodynamically favored, and the Gibbs energy of reaction is -0.65 kcal/mol, as shown in Figure 8-13. Moreover, the Gibbs energy of activation for this reaction was determined to be 1.70 kcal/mol, which is easily surmountable at room temperature. Thus, the water molecules stabilize the product state, and the calculations are in agreement with the experimental observation of facile proton transfer upon oxidation.

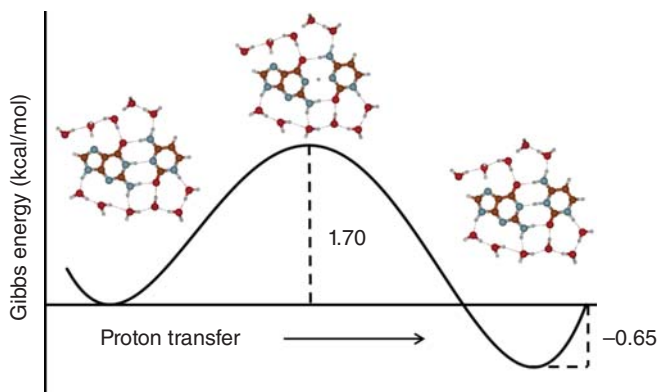


FIGURE 8-13. One-dimensional representation of the Gibbs energy surface for proton transfer in the GC radical cation base pair solvated by 11 water molecules. The Gibbs energies of the minima and TS were calculated with DFT including zero-point energy and entropic contributions. In the reactant minimum, the proton is on N1 of G, and in the product minimum, the proton is on N3 of C. At the TS, the proton is midway between N1 of G and N3 of C. Data from Ref. 5.

These examples illustrate that quantum chemistry calculations can provide useful structural, energetic, thermodynamic, and kinetic information about biological systems. As mentioned earlier, however, most biological systems are too large for a fully quantum mechanical treatment at a high level of theory. Therefore, a variety of hybrid methods that combine quantum mechanical methods with more approximate methods have been developed. In addition, conformational sampling is important for large biomolecules and is not easily included with high-level quantum chemistry methods. In particular, simple geometry optimizations will find only a very small fraction of the large number of minima and TSs on the multidimensional potential energy surfaces of biological molecules. These important issues will be discussed in more detail in Chapter 16.

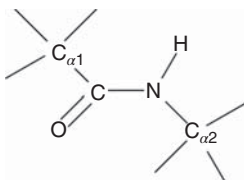
REFERENCES

The topics in this chapter are discussed in considerably more depth in a number of physical chemistry textbooks, such as References 1 – 3. The data used in the last section were obtained from References 4 and 5.

1. P. W. Atkins and J. de Paula, *Physical Chemistry*, 9th edition, Freeman, New York, 2010.
2. A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*, 1st edition, revised, McGraw-Hill, New York, 1989.
3. R. G. Parr and W. Yang, *Density-Functional Theory of Atoms and Molecules*, Oxford University Press, New York, 1989.
4. J. Sponer, J. Leszczynski, and P. Hobza, *Biopolymers* 61, 3–31 (2002).
5. A. Kumar and M. D. Sevilla, *J. Phys. Chem. B* 113, 11359–11361 (2009).

PROBLEMS

- 8-1.** Consider hydrogenic atoms, which contain one nucleus and one electron.
- What is the energy required to excite an electron from the $n = 1$ to the $n = 3$ energy level of a hydrogen atom?
 - What is the energy required to ionize the hydrogen atom?
 - How many angular and radial nodes do the $3s$, $3p$, and $3d$ orbitals have?
 - What is the degeneracy of the $n = 3$ energy level?
 - What is the energy required to ionize the He^+ atom?
- 8-2.** The bond order of a diatomic molecule is defined as $B = (N_b - N_a)/2$, where N_b is the number of electrons in bonding orbitals and N_a is the number of electrons in antibonding orbitals. The greater the bond order, the shorter the bond length and the greater the bond strength.
- By analyzing the MO diagram, arrange the diatomic molecules O_2^+ , O_2 , O_2^- , and O_2^{2-} in order of decreasing bond length.
 - How many unpaired electrons (i.e., singly-occupied MOs) does each molecule have?
 - According to Hund's rule, are the spins of the unpaired electrons parallel or antiparallel?
- 8-3.** The trihydrogen cation, H_3^+ , has been detected in the interstellar medium and is thought to be one of the most abundant ions in the universe. At the equilibrium geometry, the three hydrogen atoms are arranged in an equilateral triangle with sides 0.90 Angstroms. Here we will use MO theory to study this two-electron molecule with basis functions $1s_A$, $1s_B$, and $1s_C$, where each basis function is centered on one of the hydrogen nuclei.
- Using symmetry arguments, write the form of the lowest energy MO, $\psi_1(\mathbf{r})$, for this system. The two electrons will occupy this MO.
 - The total potential energy experienced by the nuclei is the sum of the electronic energy and the nuclear–nuclear repulsion interactions, as given in Eq. (8-9). If the Coulombic interaction between two protons separated by 1.00 Angstrom is 332 kcal/mol, what is the value of the total nuclear–nuclear interaction term in kcal/mol?
 - Will the nuclear–nuclear interaction term increase or decrease as the length of the sides of the equilateral triangle increases? Given this trend, do you expect the electronic energy to increase or decrease as the length of the sides of the equilateral triangle become longer than the value at the equilibrium geometry? Hint: Use the variation principle to answer the second part of this question.
- 8-4.** The peptide backbone of a protein is comprised of the following basic structural unit:



These six atoms lie in a plane, and this planarity plays an important role in determining the three-dimensional structures of proteins. This planarity can be understood in terms of MO theory. In this problem, we consider only the valence electrons for each atom. In other words, we do not consider the two electrons occupying the $1s$ atomic orbitals for the C, N, and O atoms. The central C atom and the N and O atoms are sp^2 hybridized, meaning that the $2s$, $2p_x$, and $2p_y$ atomic orbitals for each atom mix together to form three equivalent sp^2 atomic orbitals. The sp^2 atomic orbitals are used to form σ MOs, while the $2p_z$ atomic orbitals are used to form π MOs.

- Using symmetry arguments for the sp^2 atomic orbitals, what are the bond angles between adjacent bonds? Specifically, what are the O-C-N, C-N-H, O-C-C $_{\alpha 1}$, and C-N-C $_{\alpha 2}$ angles?
- How many valence electrons from each of these six atoms participate in σ MOs?
- How many valence electrons from each of these six atoms participate in π MOs?
- How many π MOs are formed from the three $2p_z$ atomic orbitals?
- Assuming that all of the electrons are paired, how many of these π MOs are occupied in this structural unit?

8-5. Three stationary points were identified for the proton transfer reaction in the oxidized guanine-cytosine base pair in the presence of 11 H_2O molecules, as depicted in Figure 8-13. The Gibbs energies for these three states are given in the following table with the reactant state set to zero. The values given in this table include zero-point energy and entropic contributions at 298 K.

State	Gibbs energy (kcal/mol)
Reactant: G^+-H-C	0
Product: $G-H-C^+$	-0.65
TS: $G-H^+-C$	1.70

- Is the reactant or the product thermodynamically favored?
- Calculate the equilibrium constant for this reaction at 298 K.
- Is the rate constant larger in the forward or reverse direction?
- Using transition state theory, calculate the rate constant for this reaction at 298 K.

SPECTROSCOPY

X-ray Crystallography

9.1 INTRODUCTION

The primary tool for determining the atomic structure of macromolecules is the scattering of X-ray radiation by crystals. Strictly speaking, this is not considered spectroscopy even though it involves the interaction of radiation with matter. Nevertheless, the importance of this tool in modern biology makes it mandatory to have some understanding of how macromolecular structures are determined with X-ray radiation. The basic principles of the methodology will be discussed without delving into the details of how structures are determined. We then consider some of the important results that have been obtained as well as how they can be used to understand the function of macromolecules in biological systems.

Although considerable progress has been made in determining the structures of macromolecules that are not strictly crystalline, we will concentrate on the structure of macromolecules from high-quality crystals. High-quality crystals have molecules arranged in a regular array, and this regular array, or lattice, serves as a scattering surface for X-rays.

Why are X-rays used to determine molecular structures? Ordinary objects can be easily seen with visible light, and very small objects can be seen in microscopes with a high-quality lens that converts the electromagnetic waves associated with light into an image. However, the resolution of a light microscope is limited by the wavelength of light that is used. Distances cannot be resolved that are significantly shorter than the wavelength of the light that illuminates the object. The wavelength of visible light is thousands of angstroms, whereas distances within molecules are approximately angstroms.

The obvious answer to this resolution problem is to use radiation that has a much shorter wavelength than visible light, namely X-rays, which have wavelengths in the angstrom region. In principle, all that is needed is a lens that will convert the scattering from a molecule into an image. What does this entail? We have previously shown that light has an amplitude and a periodicity with respect to time and distance, or a phase. A lens takes both the amplitude and phase information and converts it into an image. Unfortunately, a lens does not exist that will carry out this function

for X-ray radiation. Instead, what can be measured is the amplitude of the scattered radiation. Methods are then needed to obtain the phase and ultimately the molecular structure. The principles behind this methodology are given below, without presenting the underlying mathematical complications. More advanced texts should be consulted for the mathematical details (1,2).

9.2 SCATTERING OF X-RAYS BY A CRYSTAL

X-rays are produced by bombarding a target (typically a metal) with high-energy electrons. If the energy of the incoming electron is sufficiently large, it will eject an electron from an inner orbital of the target. A photon is emitted when an outer orbit electron moves into the vacated inner orbital. For typical targets, the wavelengths of the photons are tenths of angstrom to several angstroms. In a normal laboratory experiment, X-rays are produced by special electronic tubes. However, more intense short-wavelength sources can be obtained using high-energy electron accelerators (synchrotrons), and these sources are often used to obtain high-resolution structures. The intensity of X-rays from a synchrotron is thousands of times greater than a conventional source, and radiation of different wavelengths can be selected. Smaller crystals can be used for structure determinations with synchrotron radiation than with conventional sources, which is a considerable advantage.

Crystals can be considered as a regular array of molecules, or scattering elements. Only seven symmetry arrangements of crystals are possible. For example, one possibility is a simple cube. In this case, the three axes are equal in length and are 90° with respect to each other. Other lattices can be generated by having angles other than 90° and unequal sides of the geometric figure. A summary of the seven crystal systems is given in Table 9-1.

The crystal type does not give a complete picture of the possible arrangements of atoms within the crystal, that is, the lattice. The lattice is an infinite array of points (atoms) in space. A. Bravais showed that all lattices fall within 14 types, the *Bravais lattices*. For example, for a cubic crystal, three lattices are possible: one with an atom at each corner of the cube, one with an additional atom at the center of the cube (body-centered cube), and one with an additional atom in the center of each face of

TABLE 9-1. The Seven Crystal Systems

System	Axes	Angles
Cubic	$a = b = c$	$\alpha = \beta = \gamma = 90^\circ$
Rhombohedral	$a = b = c$	$\alpha = \beta = \gamma$
Tetragonal	$a = b; c$	$\alpha = \beta = \gamma = 90^\circ$
Hexagonal	$a = b; c$	$\alpha = \beta = 90^\circ; \gamma = 120^\circ$
Orthorhombic	$a; b; c$	$\alpha = \beta = \gamma = 90^\circ$
Monoclinic	$a; b; c$	$\alpha = \gamma = 90^\circ; \beta$
Triclinic	$a; b; c$	$\alpha; \beta; \gamma$

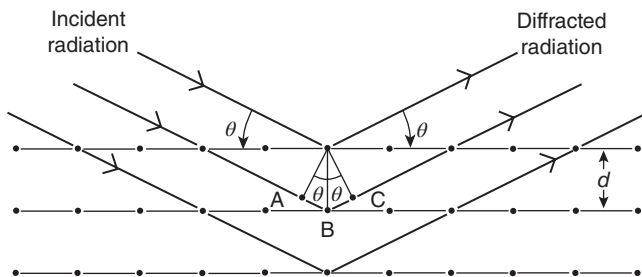


FIGURE 9-1. Diffraction of X-ray radiation by a crystal lattice. The parallel lines are planes of atoms separated by a distance, d , and the radiation impinges on the crystal at an angle θ . This diagram can be used to derive the Bragg condition for maximum constructive interference as described in the text (Eq. (9-1)).

the cube (face-centered cube). Determination of crystal type and lattice classification is important for the subsequent analysis of the X-ray scattering data.

W. L. Bragg showed that the scattering of X-rays from a crystal can be described as the scattering from parallel planes of molecules, as illustrated in Figure 9-1. If the incident beam of X-rays is at an angle θ with respect to the molecular plane, then it will be scattered at an angle θ . This is called *elastic* scattering and assumes that the radiation does not lose or absorb energy in the scattering process. In reality, some *inelastic* scattering occurs in which energy can be gained or lost, but this effect can be neglected. From Figure 9-1, it is clear that scattering will occur for each plane and scattering center with the same incident and exit angle. From a wave standpoint, the radiation will be scattered from each plane, and the radiation from each plane will have a different phase as it exits, since each wave would have traveled a different distance depending on the depth of the plane in Figure 9-1. However, if the wavelength of radiation is such that the difference in path length traveled by the beams from different planes is equal to the wavelength or an integer multiple of the wavelength, then the two waves will be in phase and constructive interference will occur, that is, the intensity of the radiation will be at a maximum. The phenomena of constructive and destructive interference have been discussed in Chapter 7. Figure 7-11 illustrates these phenomena when the time dependence of an electromagnetic wave is considered. The same analysis is applicable for the propagation of a wave in space (Fig. 7-11 and Eq. (7-20)) as illustrated in Figure 9-2.

The condition for maximum constructive interference can be calculated by reference to Figure 9-1:

$$AB + BC = n\lambda$$

where n is an integer and λ is the wavelength. The distances AB and BC are equal to $d \sin \theta$, where d is the distance between planes. Thus, the Bragg equation is

$$n\lambda = 2d \sin \theta \quad (9-1)$$

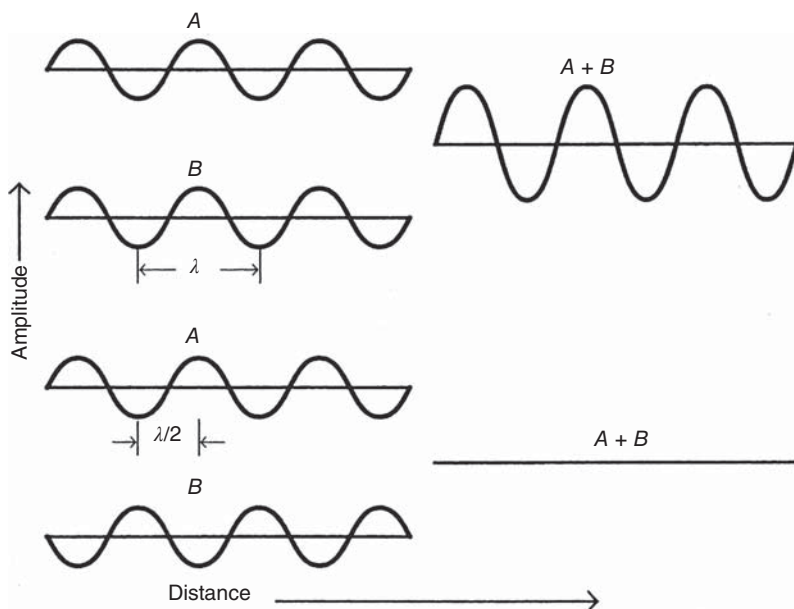


FIGURE 9-2. Illustration of constructive and destructive interference. This figure is identical to Figure 7-11, except that now the abscissa is distance rather than time. The upper two waves are in phase, which means that they differ by integral multiples of the wavelength, λ , and constructive interference occurs. The bottom two waves differ by $\lambda/2$ in phase, and destructive interference occurs.

The Bragg equation gives the condition for diffraction so that if a crystal is rotated in a beam of X-rays, the scattering pattern is a series of intensity maxima. A real crystal is more complex than a set of parallel planes of point-scattering sources. In fact, multiple planes exist, and a molecule is not a simple point scatterer. Electrons in atoms are the scatterers, and each atom has a different effectiveness as a scatterer. Consequently, when an experiment is carried out, a set of diffraction maxima of different intensities is observed. A schematic representation of the experimental setup is shown in Figure 9-3. Either the crystal or detector, or both are rotated to obtain the scattering intensity at various angles. The diffraction pattern has a strong peak at the center, the unscattered beam, and a radial distribution around the center, corresponding to different planes and values of n in Eq. (9-1).

The Bragg equation can be used to derive the minimum spacing of planes that can be resolved for X-rays of a given wavelength. Since the maximum value of $\sin \theta$ is 1, $d_{\min} = \lambda/2$.

9.3 STRUCTURE DETERMINATION

The analysis of a diffraction pattern is considerably more difficult than suggested earlier. The scattering intensity depends on the scattering effectiveness of the individual

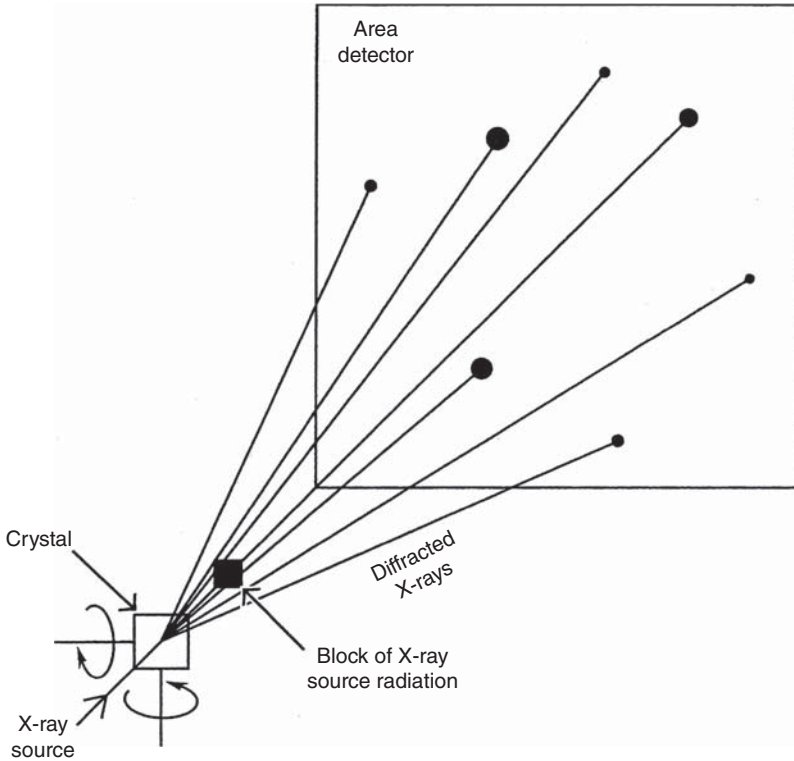


FIGURE 9-3. Experimental setup for measuring X-ray diffraction from crystals. The crystal can be rotated around both perpendicular axes, and the diffraction pattern is measured on an area detector or film. The incident X-ray radiation is prevented from hitting the area detector by insertion of a beam blocker between the crystal and the detector.

atoms and the phase of the wave from each scattering source. The structure factor, F , for each plane can be defined as the sum of the structure factors for individual atoms, f_i , times a phase factor, α_i , for each atom:

$$F = \sum_i f_i \alpha_i \quad (9-2)$$

The intensity of scattered radiation is proportional to the absolute value of the amplitude of the structure factor. The structure factor can be calculated if the atomic coordinates are known. Note that Eq. (9-2) is analogous to the general description of electromagnetic waves developed in Chapter 7, namely, an amplitude is multiplied by an angular dependence (sin and/or cos).

In proceeding to determination of the structure, it should be remembered that X-rays are scattered by electrons in atoms so that representing a crystal as a series of point atoms is not a good picture of the real situation. Instead, a more realistic

formulation is to represent the structure factor as an integral over a continuous distribution of electron density. The electron density is a function of the coordinates of the scattering centers, the atoms, and has a maximum around the position of each atom. What is desired in practice is to convert the measured structure factors into atomic coordinates. This is done by taking the Fourier transform of Eq. (9-2). In this case, the Fourier transform takes the structure factors, which are functions of the electron density, and inverts the functional dependence so that the electron density is expressed as a function of the structure factors. (This is analogous to the discussion in Chapter 7 in which the frequency and time dependences are interchanged by Fourier transforms.) This seems straightforward from the standpoint of the mathematics, but the problem is that the actual structure factors contain both amplitudes and phases. Only the amplitude, or to be more precise, its square, can be directly derived from the measured intensity of the diffracted beam. The phase factor must be determined before a structure can be calculated.

Two methods are commonly used to solve the phase problem for macromolecules. One of these is Multiple-wavelength Anomalous Dispersion or MAD. Thus far, we have discussed X-ray scattering in terms of *coherent* scattering where the frequency of the radiation is different than the frequency of oscillation of electrons in the atoms. However, X-ray frequencies are available that match the frequency of oscillation of the electrons in some atoms, giving rise to *anomalous* scattering that can be readily discerned from the wavelength dependence of the scattering. The feasibility of this approach is made possible by the use of synchrotron radiation since crystal monochromators can be used to obtain high-intensity X-rays at a variety of wavelengths. Anomalous scattering is usually observed with relatively heavy atoms (e.g., metals or iodine) that can be inserted into the macromolecular structure. The observation of anomalous scattering for specific atoms allows these atoms to be located in the structure, thus providing the phase information that is required to obtain the complete structure.

Another frequently used method for determining phases in macromolecules is *isomorphous replacement*. With this method, a few heavy atoms are put into the structure, for example, metal ions. Since scattering is proportional to the square of the atomic number, the enhanced scattering due to the heavy atoms can be easily seen, and the positions of the heavy atoms determined. This is done by looking at the difference in structure factors between the native structure and its heavy atom derivative. Of course, it is essential that the isomorphous replacement not significantly alter the structure of the molecule being studied. We will not delve into the details of the methodology here.

Thus far, the assumption has been made that the X-ray radiation is monochromatic. An alternative is to use a broad spectrum of radiation ("white" radiation). The premise is that whatever the orientation of the crystal, one of the wavelengths would satisfy the Bragg condition. In fact, this was the basis for the first X-ray diffraction experiments carried out by Max von Laue in 1912, and this approach is called Laue diffraction. As might be expected, Laue diffraction patterns can be very complex and difficult to interpret as diffraction patterns from multiple wavelengths are observed simultaneously. The advantage is that a large amount of structural information can be

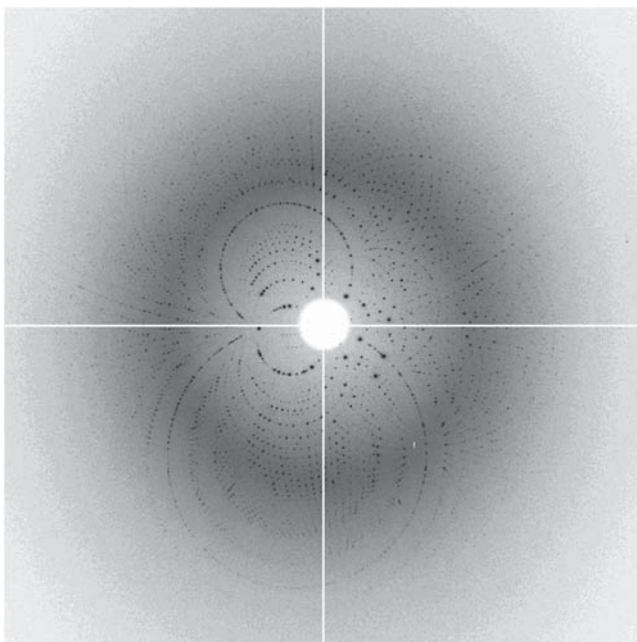


FIGURE 9-4. Laue diffraction pattern obtained with synchrotron radiation and an area detector. The protein crystal is DNA polymerase from *Bacillus stearothermophilus*. Copyright by Professor L. S. Beese, Duke University. Reproduced with permission.

obtained in a very short time. A Laue diffraction pattern of DNA polymerase from *Bacillus stearothermophilus* obtained with an area detector and synchrotron radiation is shown in Figure 9-4. Although most structural determinations use monochromatic radiation, the use of Laue diffraction has increased significantly in recent years.

What limits the precision of a structure determination? Real crystals are not infinite arrays of planes. They have imperfections so that the diffraction pattern is strong near the center and becomes weaker radially from the center, as the reflections from planes closer together become important. The precision of the distances in the final structure is limited by how many reflections are observed. The better the order in the crystal, the further from the center of the diffraction pattern that reflections can be seen. This problem can only be cured by obtaining better crystals. A minor problem is that atoms have thermal motion, so that an inherent uncertainty in the position exists. This can be helped by working at low temperatures, which sometimes will lock otherwise mobile structures into one conformation. Finally, it should be noted that in some cases, portions of the macromolecule may not be well defined because of intrinsic disorder, that is, more than one arrangement of the atoms occurs in the crystals.

The spatial resolution of a structure is usually given in terms of the distances that can be distinguished in the diffraction pattern. For good structures, this is typically 2 Å or better. However, this is not the uncertainty in the positions of atoms within the

structure. In solving the structure of a macromolecule, other information is used from known structures of small molecules, for example, bond distances (C—C, etc.) and bond angles. For high-quality structures, the uncertainty in atomic positions is tenths of angstroms.

At the present time, the primary difficulty in determining the molecular structures of proteins and nucleic acids is obtaining good crystals, that is, crystals that give good diffraction patterns. If good diffraction patterns can be obtained, including isomorphous replacements or anomalous scattering atoms, computer programs are available to help derive the structure. However, as structures become larger and larger, both obtaining good crystals and the data analysis itself are a challenge.

Crystal structures represent a static picture of the equilibrium structure so that the correlation with biological function must be approached with caution. Furthermore, the packing of molecules within a crystal can alter the structure relative to that in solution. This is usually true only for atoms on the surface of the molecule so that the core structures, such as active sites, are normally an accurate reflection of the biologically active species. Recent applications of synchrotron radiation have permitted the time evolution of structures to be studied using Laue diffraction (3,4). Although such experiments are exceedingly difficult to carry out and interpret, in principle it is possible to observe the structure of biological molecules as they function.

9.4 NEUTRON DIFFRACTION

High-velocity thermal neutrons are generated by atomic reactors. If they are slowed down by collisions, typically with D_2O , then their energy corresponds to a wavelength of about 1 Å. The scattering of neutrons by atoms is quite different than the scattering of X-rays. In the latter case, the scattering is by electrons, whereas in the former case, the scattering is by the nuclei. Moreover, the scattering of neutrons is not easily related to atomic number.

A major difference between neutron and X-ray scattering is that hydrogen is a very effective scattering center for neutrons, but is relatively ineffective for X-rays. Consequently, the positions of hydrogen atoms are difficult to obtain from X-ray scattering but can be readily found through neutron scattering. D_2O scatters quite differently than H_2O : their scattering factors have opposite signs, that is, they are out of phase with respect to each other. The scattering factors for proteins and nucleic acids usually fall somewhere in between so that mixtures of D_2O and H_2O can be used as “contrast” agents. Appropriate mixtures can be used to effectively make certain macromolecules “invisible” to neutron scattering. For example, if a protein-nucleic acid interaction is being studied, the appropriate solvent mixture can make either the protein or the nucleic acid “invisible.” Historically, this property was important for mapping the structure of the ribosome.

Neutron scattering studies are relatively rare because high-flux neutron sources are small in number. Nevertheless, neutron diffraction can be a useful tool in elucidating macromolecular structure.

9.5 NUCLEIC ACID STRUCTURE

Probably the most exciting structure determination of a biological molecule was that for B-DNA deduced by James Watson and Francis Crick. This structure was, in fact, based on diffraction patterns from fibers rather than crystals. The familiar right-handed double helix is shown in Figure 3-9. Thermodynamic aspects of this structure have been previously discussed in Chapter 3. In this structure, two polynucleotide chains with opposite orientations coil around an axis to form the double helix. The purine and pyrimidine bases from different chains hydrogen bond in the interior of the double helix, with adenine (A) hydrogen bonding to thymine (T) and guanine (G) hydrogen bonding to cytosine (C). The former pair has two hydrogen bonds and the latter has three hydrogen bonds. Hence, the G–C pair is significantly more stable than the A–T pair. Phosphate and deoxyribose are on the outside of the double helix and interact favorably with water. The double helical structure contains two obvious grooves, the major and minor grooves. The major groove is wider and deeper than the minor groove. These grooves arise because the glycosidic bonds of a base pair are not exactly opposite to each other.

We will not dwell on the biological function of DNA except to note that DNA must separate during the replication process, and the stability of various DNAs depends on the base composition of the DNA. At sufficiently high temperatures, all DNA structures are destroyed and the two polynucleotide chains separate.

Extensive X-ray studies have shown that other forms of helical DNA exist. If the relative humidity is reduced below about 75%, A-DNA forms. It is also a right-handed double helix of antiparallel strands, but a puckering of the sugar rings causes the bases to be tilted away from the normal of the axis. A-DNA is shorter and wider than B-DNA. This structure is found in biology for some double-stranded regions of RNA and in RNA–DNA hybrids.

A third type of DNA has been found, which is a left-handed helix, Z-DNA. Z-DNA is elongated relative to B-DNA and has more unfavorable electrostatic interactions. This structure has been observed with specific short oligonucleotides at high salt concentrations. The biological significance of this structure is not clear, but its occurrence demonstrates that quite different structures can exist for DNA.

RNA has more diverse structures than DNA, in keeping with its diverse biological functions that include its role as messenger RNA (mRNA) during transcription, as transfer RNA (tRNA) in protein synthesis and as ribozymes in catalysis. In addition, RNA is found in ribosomes and other ribonucleic proteins. Typically, RNA does not form a double helix from two separate chains, as DNA does. However, the same base pairing rules as found for DNA cause internal helices to be formed within an RNA molecule. In fact, the structures of many RNA molecules are inferred in this way. In RNA, uracil (U) is found rather than thymine, as in DNA.

The first structure of an RNA deduced by X-ray crystallography was that of yeast phenylalanine tRNA (5). This structure is shown in Figure 9-5. The L-shaped molecule is typical of tRNAs. The hydrogen bonding network is shown in this structure. The acceptor stem (upper right-hand corner) is where the amino acid is

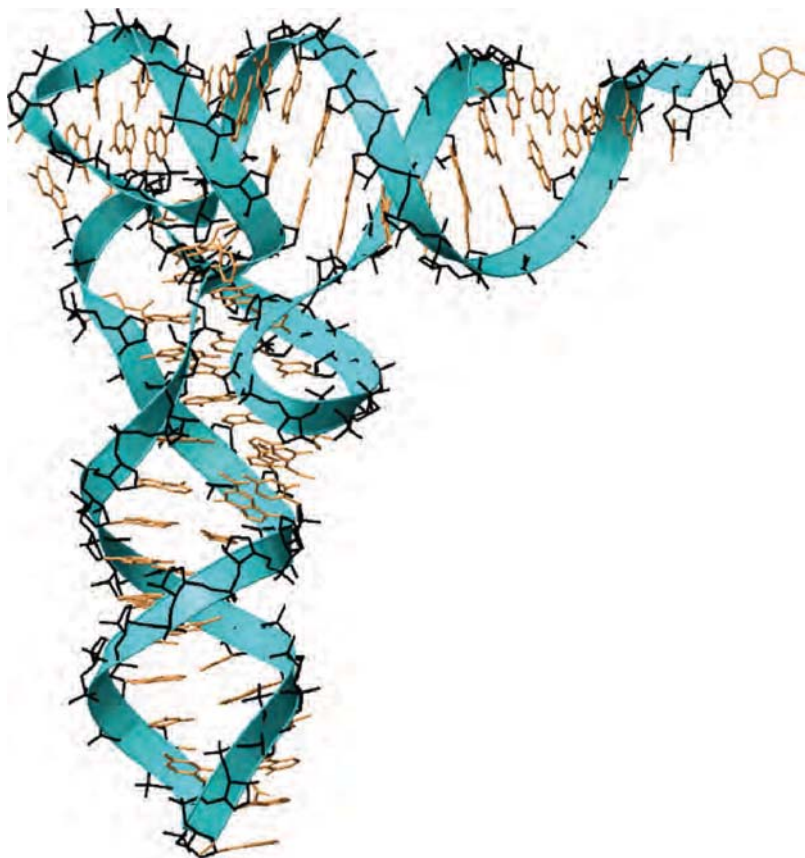


FIGURE 9-5. Structure of yeast phenylalanine tRNA with the backbone in cyan (Protein Data Bank entry 1EHZ). The L-shaped structure can be seen and the hydrogen bonding of the side chains to form helical structures. The anticodon loop is at the bottom of the long arm of the L. Copyright by Professor David C. Richardson. Reprinted with permission. Kinemage graphics, then rendered in Raster3D.

linked to form the aminoacyl-tRNA. The amino acid is transferred to the growing protein chain during protein synthesis. The anticodon, which specifies the amino acid to be added to a protein during synthesis, is at the end of the long arm of the L. It pairs with a specific mRNA that is the genetic information for the amino acid. The structures of many other tRNAs are now known and are quite similar.

For many years, a central dogma of biochemistry was that all physiological reactions are catalyzed by enzymes that are proteins. However, now many reactions are known that are catalyzed by RNA, as previously discussed in Chapter 6. In brief, these catalytic RNAs (ribozymes) are particularly important in splicing and maturation of RNA. In some cases, the ribozyme cleaves other RNA, whereas in other

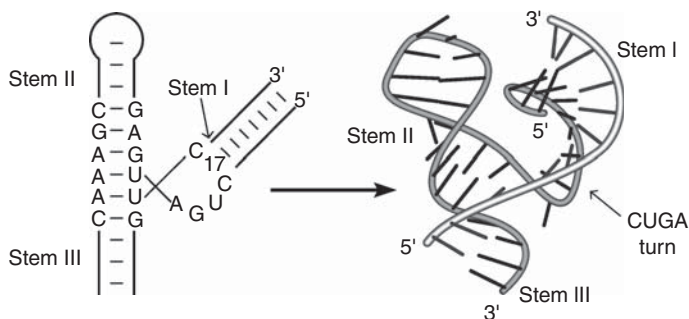


FIGURE 9-6. Secondary and tertiary structures of the hammerhead ribozyme. The capital letters refer to the bases of the nucleic acids, and the lines show hydrogen bonding interactions between the bases. The cleavage site is indicated by the arrow. Reprinted with permission from D. A. Doherty and J. A. Doudna, *Annu. Rev. Biochem.* **69**, 597 (2000). © 2000 by Annual Reviews, www.annualreviews.org.

cases it undergoes self-cleavage. Ribosomal RNA also plays a catalytic role in the formation of peptide bonds. Although there is little doubt that RNA functions as a catalyst physiologically, they are not as efficient as enzymes. In some cases, such as self-cleavage, the reaction is not truly catalytic as multiple turnovers cannot occur. Many ribozymes require a protein to function efficiently, and even in cases where true catalysis occurs, it is slow relative to typical enzymatic reactions. Kinetic aspects of ribonuclease P, a ribozyme important for the maturation of tRNA, were considered in Chapter 6.

The structures of several ribozymes have been determined (6). The first structure that was determined was that of “hammerhead” ribozyme (7). This ribozyme was first discovered as a self-cleaving RNA associated with plant viruses. The minimal structure necessary for catalysis contains three short helices and a universally conserved junction. The structure is shown in Figure 9-6, along with an indication of the cleavage site in the RNA chain. The three helices form a Y-shaped structure, with helices II and III essentially in line and helix I at a sharp angle to helix II. Distortions at the junction of the helices cause C₁₇ to stack with helix I. This structure is somewhat misleading because multiple divalent metal ions are required in order for the RNA to fold and carry out catalysis. At least one of these metal ions is intimately involved in the catalytic process. The precise number of metal ions that are required for folding and/or catalysis is uncertain. Moreover, the two structures that have been determined are necessarily noncleavable variants: one has a DNA substrate strand and the other a 2'-O-methyl group at the cleavage site. The two structures are not identical and not entirely consistent with mutagenesis studies that place specific groups at the active catalytic site. This is undoubtedly a reflection of the flexibility of the structure. Nevertheless, knowledge of ribozyme structure is a requisite for understanding their mechanisms of action and for engineering ribozymes for enhanced catalysis and function.

9.6 PROTEIN STRUCTURE

Hundreds of protein structures are known, with a great deal of variation, as well as similarities, among them. Some of the thermodynamic aspects of protein structure as well as some of the basic structures have been discussed in Chapter 3. In this section, we again return to the common motifs, as well as a few specific examples. Fortunately, the coordinates for protein structures can be found in a single database, the Protein Data Bank, which can be accessed via the Internet. Free software is also available on the Internet for viewing and manipulating structures with known coordinates, for example, Kinemage (8) and RasMol (9). This software can also be used for nucleic acid structures. Interested readers should explore this software and database, as their use is very important for understanding biological mechanisms on a molecular basis.

The linear sequence in which amino acids are arranged in a protein is termed the *primary* structure. The two most common long-range ordered structures in proteins are the α -helix and the β -sheet. Ordered structures within a protein are called *secondary* structures. Both structures are stabilized by hydrogen bonds between the NH and CO groups of the main polypeptide chain. In the case of the α -helix, depicted in Figure 3-4, a coiled rod-like structure is stabilized by hydrogen bonds between residues that are four amino acids apart. In principle, the screw sense of the helix can be right handed or left handed. However, the right-handed helices are much more stable because they avoid steric hindrance between the side chains and backbone. Helices with a different pitch, that is, stabilized by hydrogen bonding between residues other than those that are four amino acids apart are found, but the α -helix is the most common helical element found. The helical content of proteins ranges from almost 100% to very little. Helices are seldom more than about 50 Å long, but multiple helices can intertwine to give extended structures over 1000 Å long. A specific example is the interaction of myosin and tropomyosin in muscle.

Beta sheets are also stabilized by NH—CO hydrogen bonds, but in this case the hydrogen bonds are between adjacent chains, as depicted in Figure 3-5. The chains can be either in the same direction (parallel β -sheets) or in opposite directions (antiparallel β -sheets). These sheets can be relatively flat or twisted in protein structures. In schematic diagrams of proteins, α -helices are often represented as coiled ribbons and β -sheets as broad arrows pointing in the direction of the carboxyl terminus of the polypeptide chain (10).

Most proteins are compact globular structures, consisting of collections of α -helices and β -sheets. Obviously, if the structure is to be compact, it is necessary for the polypeptide chain to reverse itself. Many of these reversals are accomplished with a common structural element, a reverse turn or β -hairpin bend, that alters the hydrogen bonding pattern. In other cases, a more elaborate loop structure is used. These loops can be quite large. Although they do not have regular periodic structures, analogous to α -helices and β -sheets, they are still very rigid and well-defined structures.

The first protein whose structure was determined by X-ray crystallography to atomic resolution was myoglobin (11). Myoglobin serves as the oxygen transporter in muscle. In addition to the protein, it contains a *heme*, a protoporphyrin with a tightly

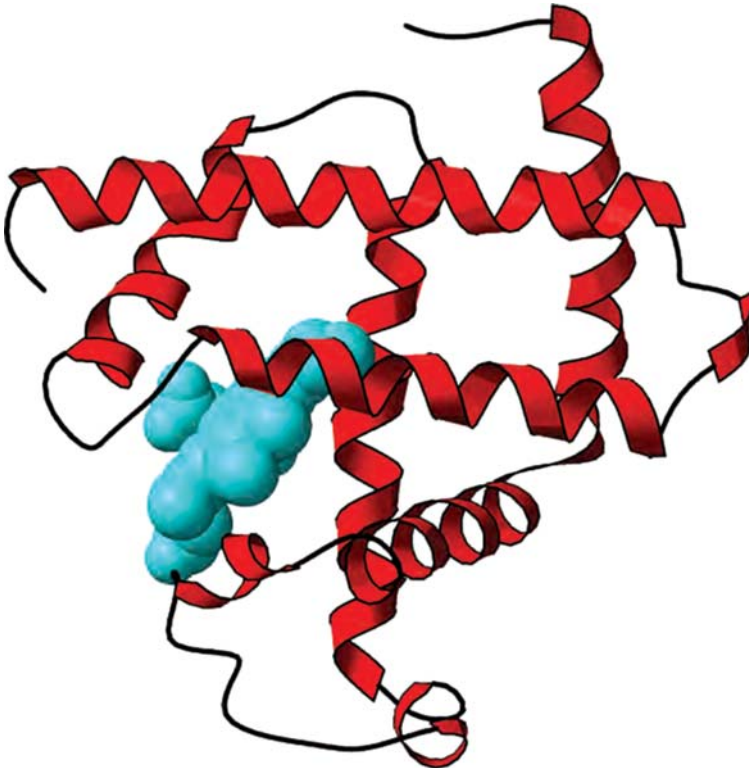


FIGURE 9-7. Ribbon structure of sperm whale myoglobin (Protein Data Bank entry 1A6M). The large amount of α -helical structure is apparent. The space filling structure in cyan is the heme. Copyright by Professor David C. Richardson. Reprinted with permission. Kinemage graphics, then rendered in Raster3D.

bound iron. The iron is the locus where oxygen is bound. Myoglobin was a particularly fortuitous choice for structure determination. Of course, good crystals could be obtained. In addition, myoglobin is very compact and primarily consists of α -helices. Its structure is shown in Figure 9-7. The α -helices form a relatively compact electron density so that the polypeptide could be easily traced at low resolution. The structure has several turns that are necessary to maintain the compact structure. Overall, the myoglobin molecule is contained within a rectangular box roughly $45 \times 35 \times 25 \text{ \AA}^3$. The overall trajectory of the polypeptide chain (that is, the folding of the secondary structure) is called the *tertiary* structure.

For water-soluble proteins, including myoglobin, the interior is primarily hydrophobic or nonpolar amino acids such as leucine, phenylalanine, etc. Amino acids that have ionizable groups, such as glutamic acid, lysine, etc., are normally on the exterior of the protein. If a protein exists in a membrane, a hydrophobic environment, the situation is often reversed, with hydrophilic residues on the inside and nonpolar residues on the outside. In some cases, however, the structure is

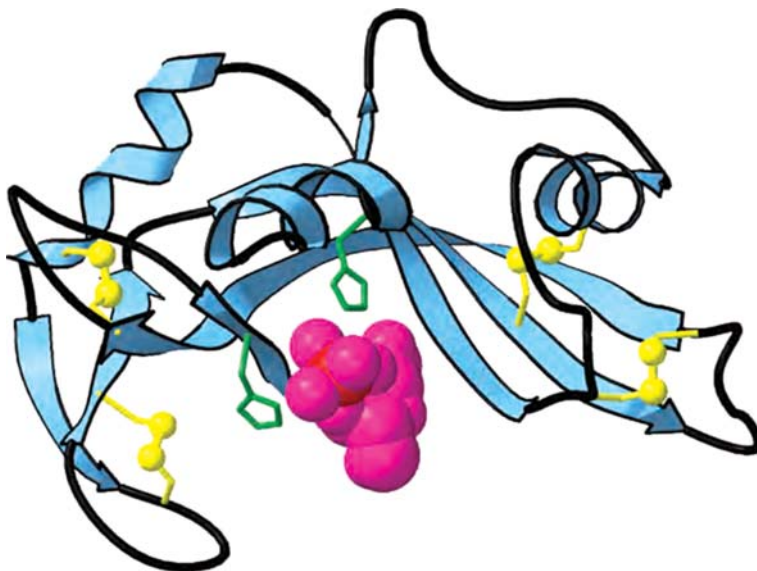


FIGURE 9-8. Schematic structure of ribonuclease A with uridine vanadate (magenta) bound to the active site (Protein Data Bank entry 1RUV). Both α -helical and β -sheet structure can be seen, as well as structure that does not have a regular array of hydrogen bonds. The disulfide linkages are in yellow, and the green residues are histidines 12 and 119, which are essential for catalytic activity. Copyright by Professor David C. Richardson. Reprinted with permission. Kinemage graphics, then rendered in Raster3D.

hydrophobic on both the “inside” and the “outside.” Understanding how polypeptides fold into their native structures is an important subject under intense investigation. In an ideal world, the structure of a protein should be completely predictable from knowledge of its primary structure, that is, its amino acid sequence. However, we have not yet reached this goal.

As a second example of a protein structure, the structure of ribonuclease A is depicted in Figure 9-8 (12). This protein is quite compact and contains both α -helix and β -sheet modules with significant loops and connecting structures. In addition, four disulfide linkages are present. This enzyme is kidney shaped with the catalytic site tucked into the center of the kidney. The active site is shown by the placement of an inhibitor and two histidine residues that participate in the catalytic reaction. This structure proved much more elusive than myoglobin because a larger variety of structural elements are present.

Not all proteins are as compact as myoglobin and ribonuclease. In some cases, a single polypeptide chain may fold into two or more separate structural domains that are linked to each other, and many proteins exist as oligomers of polypeptide chains in their biologically active form. The arrangement of oligomers within a protein is called the *quaternary* structure. The first well-studied example of the latter structure was hemoglobin. Hemoglobin is a tetramer of polypeptide chains. Two different

types of polypeptides are present: they are very similar but not identical. The typical structure of hemoglobin is $\alpha_2\beta_2$, with α and β designating the two types of quite similar polypeptide chains. Hemoglobin, of course, is used to transport oxygen in blood. In a sense, it is four myoglobins in a single molecule. Four porphyrins and irons are present, and each complex is associated with a single polypeptide chain. However, the interactions between the polypeptide chains are extremely important in the function of hemoglobin and influence how oxygen is released and taken up. In brief, the uptake and release are highly cooperative so that they occur over a very narrow range of oxygen concentration (cf. Ref. 13). This aspect of hemoglobin is discussed in Chapter 17.

Hemoglobin is known to exist in two distinct conformations that differ primarily in the subunit interactions and the details of the porphyrin binding site. The structures of both conformations of hemoglobin are shown in Figure 9-9. One of the conformations binds oxygen much better than the other, and it is the switching between these conformations that is primarily responsible for the cooperative uptake and release of oxygen. A complete discussion of hemoglobin structure is beyond the scope of this presentation, which is focused on illustrating some of the major features of protein structures.

9.7 ENZYME CATALYSIS

As a final illustration of the application of X-ray crystallography to biology, we consider the enzyme DNA polymerase. Elucidating the mechanism by which enzymes catalyze physiological reactions has been a long-standing goal of biochemistry. X-ray crystallography has been used to probe many enzyme mechanisms. The choice of DNA polymerase is arbitrary, but it is clearly an enzyme at the core of biology and requires the knowledge of both protein and nucleic acid structures. DNA polymerase is the enzyme responsible for replication of new DNA. The reaction proceeds by adding one nucleotide at a time to a growing polynucleotide chain. This addition can only occur in the presence of a *DNA template* that directs which nucleotide is to be added by forming the correct hydrogen bonds between the incoming nucleotide and the template.

The molecular structures of a wide variety of DNA polymerases are known (cf. Ref. 14). The general structure of the catalytic site is similar in all cases and has been described as a right hand with three domains, fingers, a thumb, and a palm. The structure of a catalytic fragment of the thermostable *Bacillus stearothermophilus* enzyme is shown in Figure 9-10 (15,16). The fingers and thumb wrap around the DNA and hold it in position for the catalytic reaction that occurs in the palm. To initiate the reactions, a primer DNA strand is required which has a free 3'-hydroxyl group on a nucleotide already paired to the template. The enzyme was crystallized with primer templates, which are included in Figure 9-10. The reaction occurs via a nucleophilic attack of this hydroxyl group on the α -phosphate of the incoming nucleotide, assisted by protein side chains on the protein and two metal ions, usually Mg^{2+} . We will not be concerned with the details of the chemical reaction.

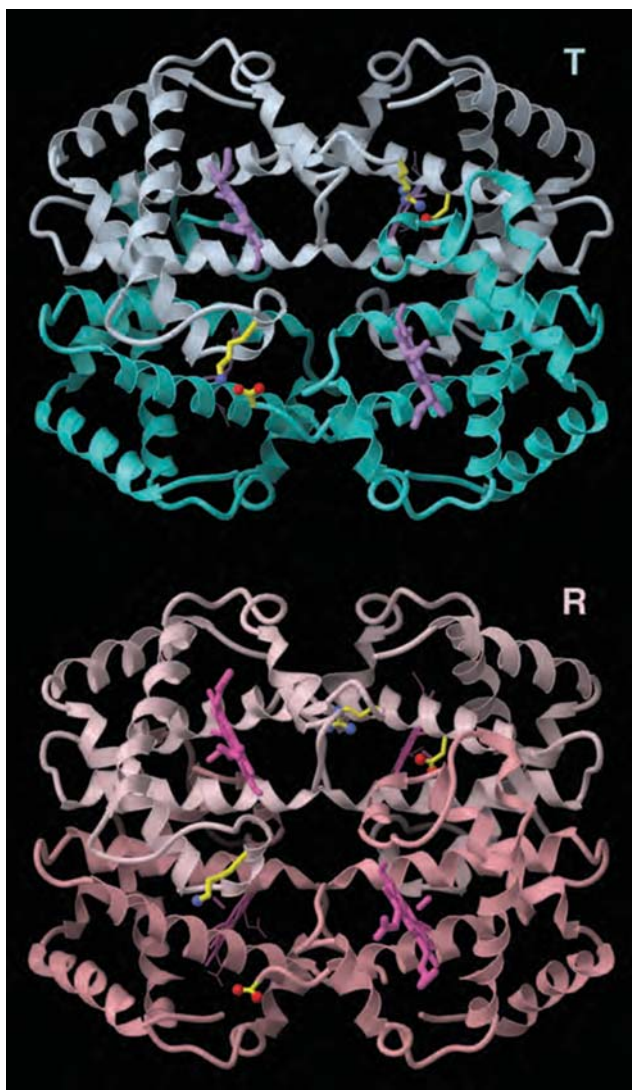


FIGURE 9-9. Schematic representation of the R (*bottom*, pink) and T (*top*, blue) forms of the hemoglobin $\alpha_2\beta_2$ tetramer. The hemes where the oxygen binds can be seen in the structure. The yellow side chains form salt bonds in the T structure that are broken in the R structure. One pair of the α - β subunits also rotates with respect to the other by about 15° in the interconversion of R and T forms. Copyright by Professor Jane Richardson. Reprinted with permission.

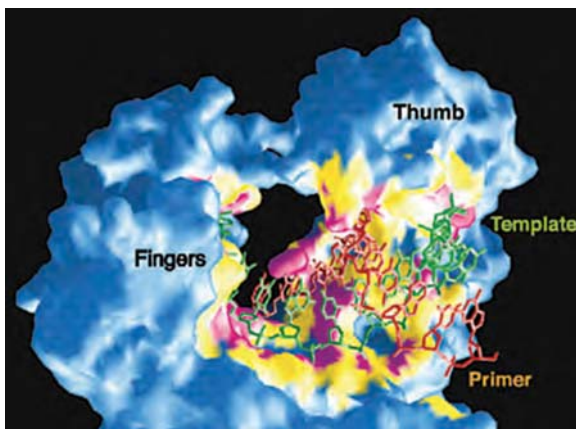


FIGURE 9-10. Structure of a catalytic fragment of DNA polymerase from *Bacillus stearothermophilus*. The “thumb” and “fingers” of the structure are labeled, and the template and primers can be seen as stick structures. Copyright by Professor L. S. Beese, Duke University. Reproduced with permission.

The DNA interacts with the protein through a very extensive network of noncovalent interactions that include hydrogen bonding, electrostatics, and direct contacts. More than 40 amino acid residues are involved that are highly conserved in all DNA polymerases. The interactions occur with the minor groove of the DNA substrate and require significant unwinding of the DNA. The major groove does not appear to interact significantly with the protein and is exposed to the solvent.

The synthesis of DNA is a highly processive reaction, that is, many nucleotides are added to the growing chain without the enzyme and growing chain separating. In the case of the polymerase under discussion, more than 100 nucleotides are added before dissociation occurs. This requires the DNA chain to be translocated through the catalytic site. This translocation has been observed to occur in the crystals by soaking the crystals with the appropriate nucleotides and determining the structures of the products of the reactions. Remarkably, the crystals are quite active, with up to six nucleotides being added and a translocation of the DNA chain of 20 Å.

The enzyme exists in multiple conformations. The initial binding of the free nucleotide occurs in an open conformation, but the catalytic step is proposed to occur in a closed conformation in which the fingers and thumb clamp onto the DNA and close around the substrates. The fidelity of DNA polymerase is remarkable, with only approximately one error per 10^5 nucleotides incorporated (17). This is accomplished primarily by the very specific hydrogen bonds occurring between the template and the incoming nucleotide, but interactions with the minor groove are also important. The conformational change accompanying binding of the nucleotide to be added probably also plays a role in the fidelity by its sensitivity to the overall shapes of the reactants. In addition, the template strand is postulated to move from a “preinsertion” site to an “insertion” site. The acceptor template base that interacts with the incoming nucleotide occupies the preinsertion site in the open

conformation and the insertion site in the closed conformation where interaction with the incoming nucleotide occurs. The structure of an “ajar” conformation occurring between the open and closed conformations, which probes incoming nucleotides for complementarity before closure around the substrate, has been determined (18). Thus, the addition of each nucleotide is accompanied by a series of conformational changes that translocate the template into position for the next step in the reaction. A movie of this process is available on the Internet (16).

Even the high fidelity of the polymerase reaction is not sufficient for the reliable duplication of genes that is required in biological systems. Consequently, the enzyme has a built-in proofreading mechanism, an exonuclease enzyme activity that is located 35 Å or more from the polymerase catalytic site. If an incorrect nucleotide is incorporated, the match within the catalytic site is not perfect, and the polymerase reaction stalls. This brief pause is sufficiently long so that the offending nucleotide base can migrate to the exonuclease site and be eliminated. Unfortunately, the proofreading process itself is not perfect so that for every 20 incorporations or so a correct base is chopped off. This wasteful process, however, increases the fidelity of replication by a factor of about 1000.

This very brief discussion of DNA polymerase indicates the great insight into biological reactions that knowledge of macromolecular structures can bring. If the crystals are biologically active, multiple structures with various substrates and inhibitors can provide a detailed mechanistic proposal. This paves the way for additional experiments that will shed even more light on the molecular mechanism. The visualization of biomolecular structures is an essential tool for the understanding of biological processes in molecular detail (19),

REFERENCES

1. B. Rupp, *Biomolecular Crystallography: Principles, Practice and Application to Structural Biology*, Garland Publishing, 2009.
2. A. McPherson, *Introduction to Macromolecular Crystallography 2nd ed.*, Wiley-Liss, Hoboken, NJ, 2009.
3. K. Moffat, *Chem. Rev.* 101, 1569 (2001).
4. R. Neutze and K. Moffat, *Curr. Opinions Struct. Biol.* 32, 651 (2012).
5. A. Rich and S. H. Kim, *Sci. Am.* 238, 52 (1978).
6. E. A. Doherty and J. A. Doudna, *Annu. Rev. Biochem.* 69, 597 (2000).
7. W. G. Scott, J. T. Finch, and A. Klug, *Cell* 81, 991 (1995).
8. D. C. Richardson and J. S. Richardson, *Trends Biochem. Sci.* 19, 135 (1994).
9. R. A. Sayle and E. J. Milner-White, *Trends Biochem. Sci.* 20, 374 (1995).
10. J. S. Richardson, D. C. Richardson, N. B. Tweedy, K. M. Gernert, T. P. Quinn, M. H. Hecht, B. W. Erickson, Y. Yan, R. D. McClain, M. E. Donlan, and M. C. Suries, *Biophys. J.* 63, 1186 (1992).
11. C. L. Nobbs, H. C. Watson, and J. C. Kendrew, *Nature* 209, 339 (1966).
12. G. Kartha, J. Bello, and D. Harker, *Nature* 213, 862 (1967).
13. M. F. Perutz, A. J. Wilkinson, M. Paoli, and G. G. Dodson, *Annu. Rev. Biophys. Biomol. Struct.* 27, 1 (1998).

14. T. A. Steitz, *J. Biol. Chem.* 274, 17395 (1999).
15. J. R. Kiefer, C. Mao, J. C. Braman, and L. S. Beese, *Nature* 391, 304 (1998).
16. S. J. Johnson, J. S. Taylor, and L. S. Beese, *Proc. Natl. Acad. Sci. USA* 100, 3895 (2003).
17. T. A. Kunkel and K. Bebenek, *Annu. Rev. Biochem.* 69, 497 (2000).
18. E. Y. Wu and L. S. Beese, *J. Biol. Chem.* 256, 19758 (2011).
19. Several Internet sites are available for the visualization of structures generated from data bases of molecular coordinates: www.proteinexplorer.org, <http://kinemage.biochem.duke.edu>, and <http://rasmol.org/>.

PROBLEMS

- 9-1. Assume a lattice of atoms equidistant from each other in all directions (a cubic lattice) with a distance between atoms of 2.86 Å. If a crystal of this material is irradiated with X-rays with a wavelength of 0.585 Å, at what angles are Bragg reflections seen for the planes that are 2.86 Å apart? (More than one set of reflections will be seen, but we will not deal with that complexity here.) If the X-rays have a wavelength of 6.00 Å, what would be observed?
- 9-2. The bond energy for a C—C bond is about 340 kJ/mol, 600 kJ/mol for C=C, and 400 kJ/mol for C—H. Calculate the energy/mol associated with X-ray radiation with wavelengths of 0.585 and 1.54 Å. What does this imply about the stability of biomolecules in an X-ray beam?
- 9-3. When a molybdenum target is used as an X-ray source, X-rays of wavelength 0.710 Å are produced. For a particular crystal, a reflection was observed at 4° 48'. What are the possible distances between Bragg planes? The same reflection observed when a copper target is used as an X-ray source is found at 10° 27'. What is the wavelength of the X-ray produced by the copper target?
- 9-4. A crystal structure consists of identical units called unit cells. The unit cells are arranged in order, much like a brick wall, to form the crystal. The unit cell contains multiple atoms, and often multiple molecules. Consider one particular plane of a unit cell that extends indefinitely in the X direction. The following structure factors, F_i , were obtained for the planes (Data from P. W. Atkins, *Physical Chemistry*, 3rd edition, Freeman, New York, 1986, p. 561.):

Plane	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
F_i	16	-10	2	-1	7	-10	8	-3	2	-3	6	-5	3	-2	2	-3

Assume that the electron density, $\rho(x)$, can be calculated from the structure factors by the relationship

$$\rho(x) = F_0 + 2 \sum_{i>0} F_i \cos(2\pi ix)$$

Evaluate this sum for $x = 0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9$, and 1.0 , and construct a plot of $\rho(x)$ versus x . (This is best done with a computer program.) The maxima in the plot correspond to positions of atoms along the x -axis. This exercise should give you a good idea of how electron density maps are obtained. The equation for $\rho(x)$ is an example of a Fourier transformation.

Electronic Spectra

10.1 INTRODUCTION

The most common type of spectroscopy involves light in the visible and ultraviolet regions of the spectrum interacting with molecules. This interaction causes electrons to shift between their allowed energy levels. If light shines on a colored sample, some of the radiation is absorbed by the sample. This absorption can heat the sample, it can cause photons of the same or lower energy to be emitted, or photochemical reactions can occur. We will not be concerned with the latter phenomenon. Of course, in all cases, energy must be conserved, that is, the total energy of all of the three processes must be equal to the energy of the photons that are absorbed.

The absorption of light is due to the interaction of the oscillating electromagnetic field of the radiation with the charged particles in the molecule. If the electromagnetic force is sufficiently large, the electrons in the molecule are rearranged, that is, shifted to higher energy levels. The absorption process is extremely fast, occurring in about 10^{-15} s. The excited state equilibrates with its surroundings and returns to the ground electronic state either by the production of heat or by the emission of radiation with an energy equal to or less than that of the absorbed radiation. The latter process is called *fluorescence* and typically occurs in nanoseconds. The absorption process is shown schematically in Figure 10-1; a more exact description is discussed later (Fig. 10-11).

The details of light absorption by a molecule can be determined directly from quantum mechanical calculations. A transition of electrons from one energy level to another occurs when the two electronic states in question are coupled. The detailed calculation tells us what transitions between the energy levels are allowed, for example, from energy level 1 to energy level 2. Electrons can only move between certain energy states; not all transitions are allowed. These *selection rules* are important because they allow us to predict what wavelengths of light will be absorbed and emitted. Although this facet of spectroscopy will not be of great concern here, it is important to remember that transitions of electrons between energy levels are governed by a set of rules. Exceptions to selection rules sometimes occur, but with a much lower probability than the favored event. This is primarily

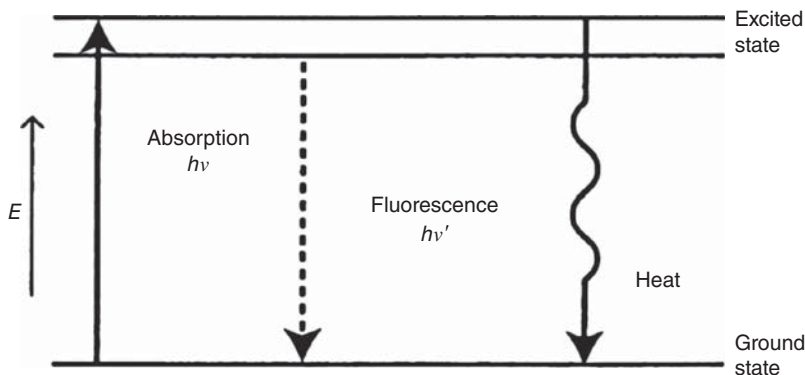


FIGURE 10-1. Simplified schematic diagram of electron excitation by absorbance of radiation. An electron is excited to a higher energy level by radiation of frequency ν . It returns to its ground state through fluorescence at a frequency ν' and/or by dissipation of heat. A more complete diagram is presented in Figure 10-11.

because the potential energy functions used in the quantum mechanical calculations are not perfect.

10.2 ABSORPTION SPECTRA

From a practical standpoint, the most important aspect of the quantum mechanical calculation is the determination of how much light is absorbed by the sample. This is embodied in the Beer–Lambert law, which gives the relationship between the light intensity entering the solution, I_0 , and the light intensity leaving the solution, I :

$$\log \left(\frac{I_0}{I} \right) = A = \epsilon cl \quad (10-1)$$

Here A is defined as the absorbance, ϵ is the molar absorptivity or extinction coefficient, c is the concentration of the absorbing material, and l is the thickness of the sample through which the light passes. In visible–ultraviolet spectrophotometers, a solution of the sample is put into a cuvette of calibrated dimensions, and visible or ultraviolet light is passed through the sample. The intensity of the light before and after passing through the solution is determined through the use of a variety of detectors, such as photomultipliers and photodiodes. A schematic diagram of a typical spectrophotometer is shown in Figure 10-2. The extinction coefficient is a different constant for each wavelength and is characteristic of the molecule under investigation. In principle, it can be calculated through quantum mechanics, but more practically it can be determined experimentally. A spectrophotometer also must account for light lost by reflection, absorption by the cuvette, and various other factors.

Deviations from the Beer–Lambert law can be observed in the laboratory. For example, the light might not be monochromatic, the sample might be inhomogeneous,

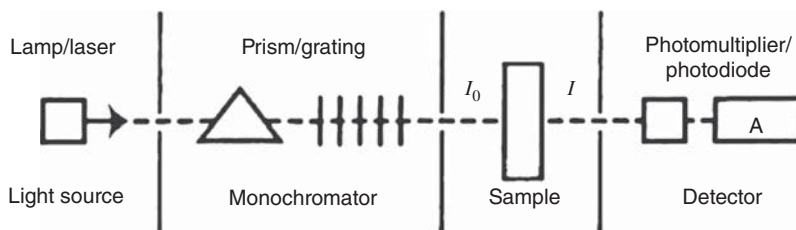


FIGURE 10-2. Schematic representation of an absorption spectrophotometer. Common light sources are tungsten and arc lamps and lasers. Monochromatic light of intensity I_0 is obtained with a prism or diffraction grating. The detection of the transmitted light of intensity I is with a photomultiplier or photodiode, with the electronic signal converted to the absorbance, A .

light scattering could occur, the photodetectors might exhibit nonlinear behavior, etc. These are instrumental and sample artifacts. However, apparent deviations that are due to the molecular properties of the sample also occur: for example, sample aggregation or complex formation as the concentration is changed. In this case, the deviations can be used to study these properties. Experimental studies must carefully verify the validity of the Beer–Lambert law over the concentration range being studied.

A very important property of light absorbing solutions is that if multiple absorbing species are present, the total absorbance is simply the sum of the absorbance of the individual species. For example, for species X and Y , the absorbance is

$$A_\lambda = A_\lambda^X + A_\lambda^Y = \epsilon_\lambda^X l[X] + \epsilon_\lambda^Y l[Y] \quad (10-2)$$

If measurements are made at two different wavelengths where the extinction coefficients are sufficiently different, it is possible to determine the concentrations of the individual species if the extinction coefficients are known. Equation (10-2) generates two equations with different numerical values of the extinction coefficients at the two wavelengths. These two equations can be solved simultaneously to give the two concentrations. This is a very common method for determining the concentrations of individual species in solutions. In fact, this approach can be extended to more complicated situations. For example, if three species are present, absorbance measurements at three different wavelengths can be used to determine the concentrations of the individual components.

The wavelength at which two components have exactly the same extinction coefficient is called the *isobestic* wavelength. In this case, it is obvious that measurement of the absorbance determines the total concentration of the two components:

$$A_{\text{isobestic}} = \epsilon_{\text{isobestic}} l([X] + [Y]) \quad (10-3)$$

Although one or more isobestic wavelengths may be found for any pair of components, this is not necessarily the case, as the two components may not have significant absorbance in the same wavelength region. A useful rule to remember is that if two

or more isobestic wavelengths are seen at a series of different concentrations and/or varying conditions such as pH, then only two components are present. (This is not entirely rigorous, but the probability of this not being the case is exceedingly small.) This is especially useful when examining samples in which the individual extinction coefficients are not known.

10.3 ULTRAVIOLET SPECTRA OF PROTEINS

Proteins that do not contain strongly absorbing extrinsic cofactors such as hemes have a very characteristic ultraviolet spectrum, whereas they are essentially transparent in the visible region of the spectrum. The spectrum of serum albumin is shown in Figure 10-3 as a plot of the absorbance versus the wavelength. It has two absorption maxima: one at about 280 nm and the other, much stronger absorption, at around 190–210 nm. The absorption at 280 nm is frequently used to measure protein concentrations. Although the absorption is much stronger at 200 nm, it is harder to make measurements at this wavelength, so absorption at 280 nm is not routinely used as an analytical tool. The ultraviolet spectrum of a protein is sufficiently characteristic to use as a preliminary indication that a protein is present.

The spectra of the aromatic amino acids phenylalanine, tyrosine, and tryptophan are shown in Figure 10-4. It can be readily seen that electronic transitions in tyrosine

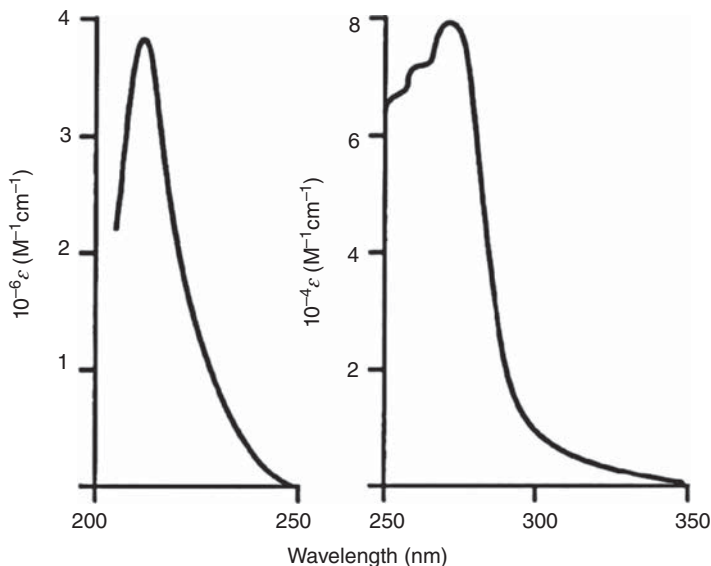


FIGURE 10-3. Ultraviolet absorption spectrum of bovine serum albumin at pH 7.0. The extinction coefficient is plotted versus the wavelength in two panels because of the large difference in the ordinate scale as the wavelength changes. Data from E. Yang in I. Tinoco, K. Sauer, J. C. Wang, and J. D. Puglisi, *Physical Chemistry*, 4th edition, Prentice-Hall, Englewood Cliffs, NJ, 2002, p. 548.

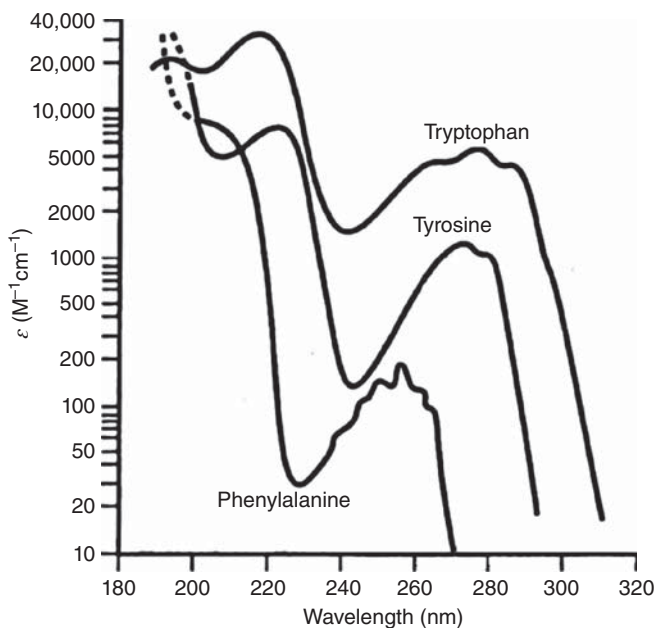


FIGURE 10-4. Absorption spectra of phenylalanine, tyrosine, and tryptophan at pH 6.0. Note the logarithmic scale on the ordinate. Adapted from S. Malik in D. B. Wetlaufer, *Ultraviolet spectra of proteins and amino acids*, *Adv. Protein Chem.* **17**, 303 (1962). © 1962, with permission from Elsevier.

and tryptophan are primarily responsible for the absorption peak at around 280 nm. It can also be seen that all three amino acids absorb strongly at shorter wavelengths. Measurements of the spectra of polypeptides not containing aromatic amino acids show strong absorbance at around 192 nm due to electronic transitions associated with the peptide bond. Thus, the absorbance maximum in the 190–210 nm range is due to both aromatic amino acids and peptide bonds. The absorption due to side chains of other amino acids is typically at least an order of magnitude less in this wavelength region.

The spectra of aromatic amino acids and peptide bonds in proteins are significantly influenced by their local environments. This is because the spectra are fundamentally due to electronic phenomena so that structural features that determine the local charge distribution and the local dielectric constant strongly influence the spectra. Thus, the spectra of peptide bonds and aromatic amino acids buried inside the protein are somewhat different than those on the exterior of the protein. Hydrogen bonding and ion pair interactions also influence the spectra. This environmental sensitivity of protein spectra can be used to obtain information about the structure of proteins. For example, the amount of α -helix, β -sheet, and random coil can be estimated from the ultraviolet spectrum (1). A denatured protein usually has a somewhat different ultraviolet spectrum than the native protein so that protein folding reactions can be monitored by measuring spectral changes. However, the change in the ultraviolet spectrum is

usually rather small so that other methods have proved more useful for assessing secondary structure and monitoring protein folding (e.g., circular dichroism, which is discussed later).

10.4 NUCLEIC ACID SPECTRA

In contrast to proteins, all of the common nucleotide bases have quite similar absorption spectra, with an absorption peak at around 260 nm. The spectra of adenine, thymine, cytosine, uracil, and guanine are shown in Figure 10-5. Nucleic acid solutions also are essentially transparent in the visible region of the spectrum. The intense absorption at 260 nm is frequently used as an indication of the presence of nucleic acids and can be used to determine their concentrations by use of the Beer–Lambert law.

Also, the local environment of individual nucleotide bases can influence the spectrum. Generally, nucleic acids have a lower absorbance than the sum of the absorbance of the individual nucleotides. This phenomenon is called *hypochromicity* and is due to the “stacking” or lining up of parallel planes of the bases. An increase in absorption is called *hyperchromicity*. These phenomena are very useful for determining whether nucleic acids are structured. For example, native (double-helix) DNA displays significant hypochromicity, and if the DNA is unwound, the absorbance at 260 nm increases.

Although the stable structure of DNA is the well-known double helix, the double-helix structure is lost as the temperature increases, and at a sufficiently high temperature, the two strands of DNA separate. This loss of structure can be monitored by the increase in absorbance at 260 nm, as shown schematically in Figure 10-6. Obviously, the more stable the DNA, the higher the temperature that is needed to break down the structure so that doing a temperature “melt” of DNA is a useful tool for determining the stability of a specific DNA. With this technique, for

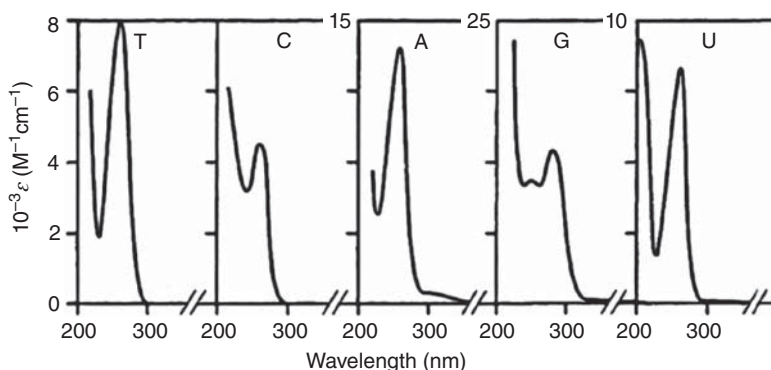


FIGURE 10-5. Absorption spectra in water of the four bases commonly found in nucleic acids: thymine (T), cytosine (C), adenine (A), guanine (G), and uracil (U). Data from H. Du, R. A. Fuh, J. Li, A. Corkan, and J. S. Lindsey, *Photochem. Photobiol.* **68**, 141 (1998) [<http://omlc.ogi.edu/spectra/PhotochemCAD>].

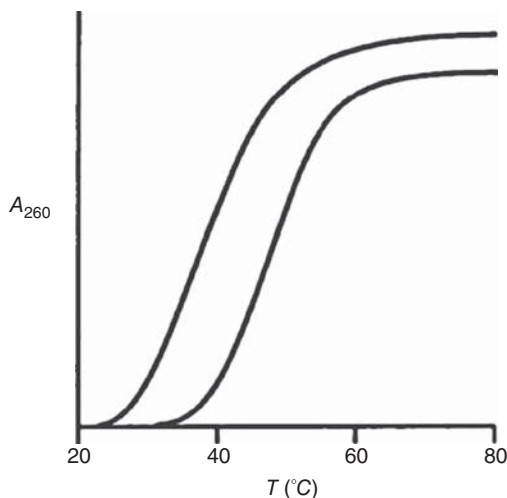


FIGURE 10-6. Schematic representation of DNA melting for two different DNA. The absorbance at 260 nm increases as the temperature is raised and the double-stranded structure breaks down. The temperature at which the melting occurs increases as the amount of G–C base pairs increases.

example, it is easy to show that G–C base pairs are more stable than A–T pairs. By studying the stability of a series of nucleic acid oligomers, it is possible to predict the stability of DNA and, less reliably, RNA structures (cf. Chapter 3). Also, when DNA is synthesized, the double-helix structure must be broken in order to provide a single-strand template for the polymerization reaction. The kinetics of DNA structure formation/ breakdown has been discussed in Chapter 6. The unwinding of DNA *in vivo* is done by helicase enzymes. Obviously, monitoring the DNA structure is important for understanding DNA synthesis, in general, and the mechanism of action of helicases, in particular. This subject is considered further later in this chapter.

10.5 PROSTHETIC GROUPS

Some proteins contain tightly bound organic molecules and/or metal ions. These are usually referred to as *prosthetic groups* and are essential for the biological activity of the macromolecules. The spectra of these prosthetic groups are often in the visible region of the spectrum, easily separable from the absorption due to the protein, and can be monitored to follow the reactions undergone by the protein.

One of the most common prosthetic groups is *heme*, the structure of which is shown in Figure 10-7. The heme binds iron tightly, either as Fe^{2+} or as Fe^{3+} . Probably, the most well-known heme protein is hemoglobin, mentioned in Chapter 9. Hemoglobin has a molecular weight of about 64,000. It has four polypeptide chains and contains four hemes. It is, of course, responsible for the transport of oxygen in

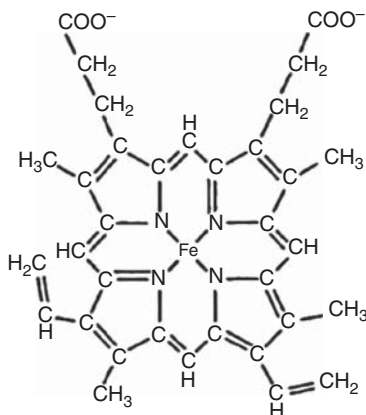


FIGURE 10-7. Structure of the heme found in hemoglobin. The Fe is in the oxidation state +2 when oxygen is bound.

the body. The structure of the protein is well known (Fig. 9-9) and its function has been extensively studied (cf. Ref. 2). For the purposes of this discussion, we consider only the visible spectra of the molecule. The active molecule contains Fe^{2+} : when oxygen is bound, it is coordinated directly to the iron. In the absence of oxygen, this iron coordination site is occupied by water. The visible spectra of oxy- and deoxyhemoglobin are quite distinct, as shown in Figure 10-8. This is because the

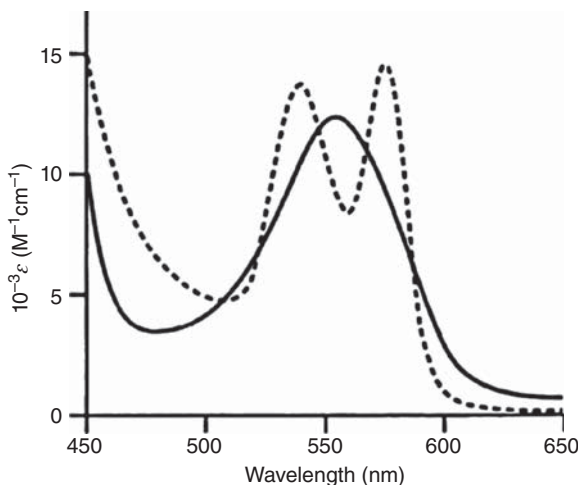


FIGURE 10-8. Visible absorption spectra of oxyhemoglobin (dashed line) and deoxyhemoglobin (solid line). These two forms of hemoglobin can be readily distinguished with absorption spectroscopy. Adapted from L. Stryer, *Biochemistry*, 2nd edition, Freeman, San Francisco, CA, 1981, p. 52. © 1976, 1981 by Lubert Stryer. Used with permission of Freeman.

binding of oxygen alters the environment of the iron/heme, thereby altering the electronic energy levels. Thus, the visible spectrum can be used to monitor the binding of oxygen. Oxyhemoglobin is responsible for the bright red color of oxygenated blood. The ferrous iron can be oxidized to the ferric form, which also alters the spectrum of the molecule. This form of hemoglobin is not involved in oxygen transport. The molecular mechanism for the binding of oxygen to hemoglobin is now understood in molecular detail. It is a highly cooperative process with alterations in the position of the iron in the molecule and significant conformational changes within the protein accompanying binding (cf. Chapter 17).

Cytochromes also contain hemes. For these molecules, the oxidation-reduction of iron is crucial to their physiological function. The progress of the oxidation-reduction reactions can be readily monitored through changes in the visible spectrum.

Flavins constitute another type of common prosthetic group. Flavin mononucleotide (FMN) and flavin adenine dinucleotide (FAD) are important constituents in many enzymes that participate in redox reactions. Their structures are shown in Figure 10-9. These flavin prosthetic groups can exist in three oxidation states: oxidized, semiquinone radical, and reduced. (A radical is an organic molecule with an unpaired electron.) All three of these oxidation states have different absorption peaks in the visible. The three oxidation states and their absorption maxima are included in Figure 10-9. The semiquinone can also lose a hydrogen to give a diradical (two unpaired electrons) with yet a different absorption maximum. The pK for the loss of this hydrogen is about 8.4. Flavin enzymes can undergo both one- and two-electron oxidation-reduction reactions and can be readily reoxidized by oxygen. The visible spectrum serves as a useful indicator of the oxidation state of the flavin in an enzyme. It can also be used to monitor the reaction progress of the enzymatic reactions.

Why do these prosthetic groups have electronic transitions at longer wavelengths than proteins? This can be readily understood in molecular and quantum mechanical terms. Both hemes and flavins have highly aromatic structures, that is, extensive conjugation occurs. As a result, the electrons are delocalized around the ring structures (π electrons). If the electrons are considered as “particles in a box,” as discussed in Chapter 7, the size of the box is increased relative to less conjugated structures such as tyrosine, phenylalanine, and tryptophan. An increase in the size of the box increases the wavelengths that characterize the electronic transitions (Eq. (7-12)).

10.6 DIFFERENCE SPECTROSCOPY

Because the absorption of light by proteins and nucleic acids depends on the local environment of the chromophores, changes in light absorption often occur when small molecules bind to proteins or nucleic acids and when macromolecules interact. As an example, we consider the binding of nucleotides to the enzyme ribonuclease A. As the name implies, ribonuclease A hydrolyzes RNA. It is a small protein with a molecular weight of approximately 14,000 and contains tyrosine, tryptophan, and phenylalanine so that it has a typical protein ultraviolet spectrum. Its structure is shown in Figure 9-8.

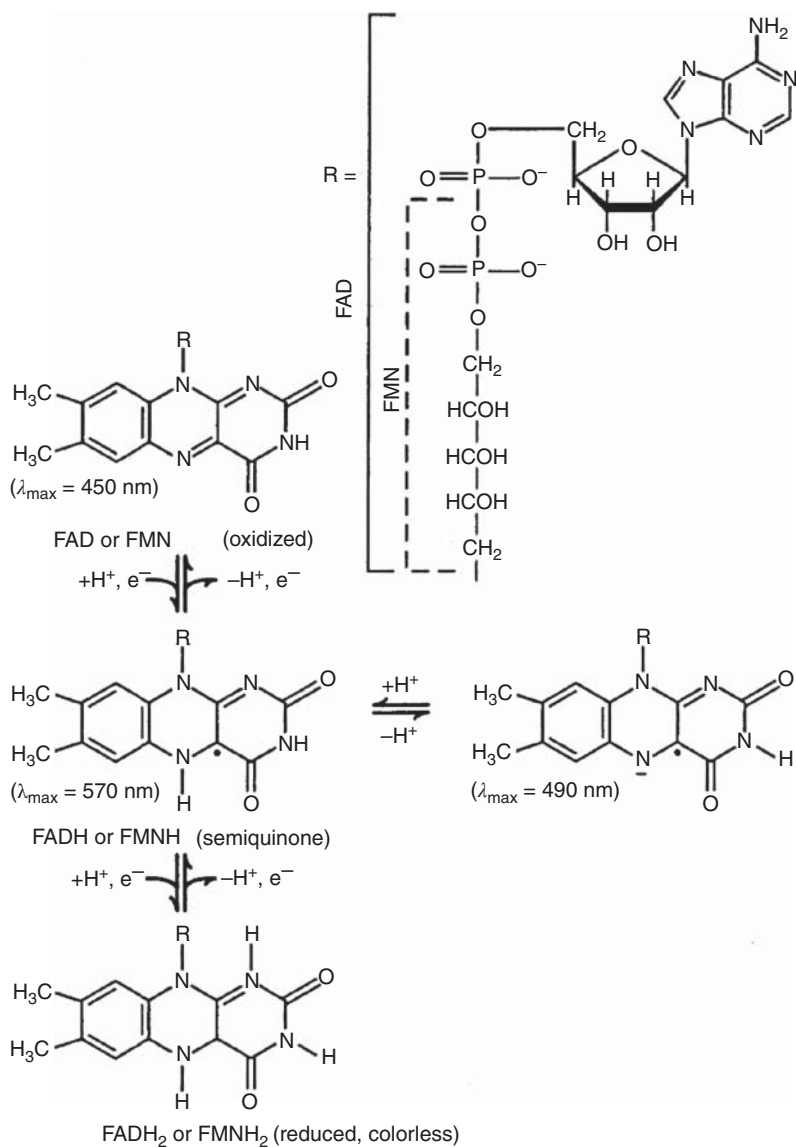


FIGURE 10-9. Oxidation states of FAD and FMN with the absorption maxima indicated. These oxidation states are important for the biological function of these molecules and can be readily distinguished with absorption spectroscopy. The pK of the semiquinone is about 8.4.

It binds nucleic acids and has a great preference for cleavage adjacent to a pyrimidine. When nucleotides bind to the enzyme, the ultraviolet spectrum is changed. However, the changes are very small so that a technique called difference spectroscopy must be used.

With difference spectroscopy, a double-beam spectrophotometer is used. Each beam contains two cuvettes. For one of the beams, one cuvette contains the nucleotide and the other the protein. In the other beam, one of the cuvettes contains a mixture of the nucleotide and the protein and the other contains buffer. The concentrations of nucleotide and the enzyme are exactly the same in each beam. Consequently, any difference in absorbance that is measured must be due to the interaction of the nucleotide and the protein. The absorbance in the beam containing the mixture of enzyme and nucleotide (assuming a path length of 1 cm) is

$$\begin{aligned} A_1 &= \epsilon_{EL}[EL] + \epsilon_E[E] + \epsilon_L[L] \\ &= \epsilon_{EL}[EL] + \epsilon_E([E_0] - [EL]) + \epsilon_L([L_0] - [EL]) \end{aligned} \quad (10-4)$$

where ϵ 's are the extinction coefficients for the enzyme, E, the ligand, L, and the enzyme–ligand complex, EL. The study of ligand binding showed that only one nucleotide binds per enzyme molecule. The absorbance in the other beam is given by

$$A_2 = \epsilon_E[E_0] + \epsilon_L[L_0] \quad (10-5)$$

Subtraction of Eq. (10-5) from Eq. (10-4) gives the difference absorbance

$$\Delta A = (\epsilon_{EL} - \epsilon_E - \epsilon_L)[EL] = \Delta\epsilon[EL] \quad (10-6)$$

Thus, measuring the difference absorbance is a direct measure of the concentration of the enzyme–nucleotide complex.

Titration of the enzyme with the nucleotide and extrapolation to a concentration where all of the ligand is bound to the enzyme permits the determination of the difference extinction coefficient, $\Delta\epsilon$. The difference extinction coefficient for the binding of 2'-cytidine monophosphate to ribonuclease is shown in Figure 10-10 at different wavelengths. Note that significant changes are found at wavelengths of 260 and 280 nm, indicating that the extinction coefficients of both the enzyme and the nucleotide are altered in the complex. Note also the isobestic point that is found as the pH is varied, demonstrating the presence of only two absorbing species, bound and unbound. Since the concentration of all species can be determined during the titration, the binding constant can be calculated and was found to be $1.8 \times 10^5 \text{ M}^{-1}$ at pH 5.5. An extensive study carried out with various ligands over a range of pH provided valuable information about the molecular details of the interaction between the ligands and enzyme (3).

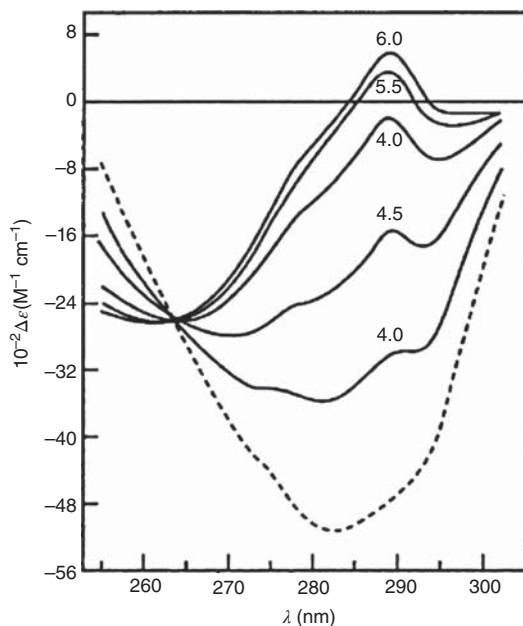


FIGURE 10-10. Difference absorption spectra of ribonuclease A-2'-CMP as a function of pH. The pH is indicated next to each difference spectrum. The dashed line is the calculated difference spectra for the protonated ring species of 2'-CMP. Reprinted in part with permission from D. G. Anderson, G. G. Hammes, and F. G. Walz, Jr., *Biochemistry* 7, 1637 (1968). © 1968 by American Chemical Society.

10.7 X-RAY ABSORPTION SPECTROSCOPY

X-rays can also be absorbed by materials, and the Beer–Lambert law (Eq. (10-1)) is applicable. The energy associated with an X-ray of 1 Å wavelength is about 10^6 kJ/mol ($E = hc/\lambda$). This energy is sufficiently large that the absorption of X-rays by atoms results in electrons being ejected from inner orbits (1s, 2s, 2p). The X-ray absorption spectra of atoms consist of sharp peaks, the peaks occurring at different wavelengths for every element in the periodic table. If the atom is in a molecule, the ejected electron interacts with the environment, resulting in small oscillations around the main peak in the absorption spectra due to back scattering. Analysis of this fine structure provides information about the nature of the atoms surrounding the atom that has absorbed the X-rays. This is particularly useful for probing the environment around heavy metals in biological structures.

10.8 FLUORESCENCE AND PHOSPHORESCENCE

Fluorescence is a very common phenomenon in biology and elsewhere. As previously discussed, fluorescence is the emission of light associated with electrons moving from

an excited state to the ground state. Fluorescence and phosphorescence are often more useful tools for studying biological processes than absorbance. They can be considerably more sensitive than absorbance so that much lower concentrations can be detected. They are also often extremely sensitive to the environment and thus are extremely good probes of the local environment. Some other useful properties will be discussed a bit later.

In order to understand fluorescence and phosphorescence, a somewhat more complex energy diagram such as Figure 10-11 is useful. In this diagram, two electronic energy levels are shown, each with a manifold of vibrational energy levels. The vibrational energy levels are much more closely spaced than the electronic energy levels and are due to the vibrations of atoms within a molecule. Different modes of vibration have different energies so that manifolds of quantized energy levels exist. The properties of vibrational energy levels can be calculated through quantum mechanics. Generally, electrons are in their ground electronic energy state at room temperature, and the molecules are also in their lowest vibrational energy level. They can be excited by light to the next electronic energy level, but can be in many different vibrational energy levels in the excited state. After this excitation, the molecule will return to

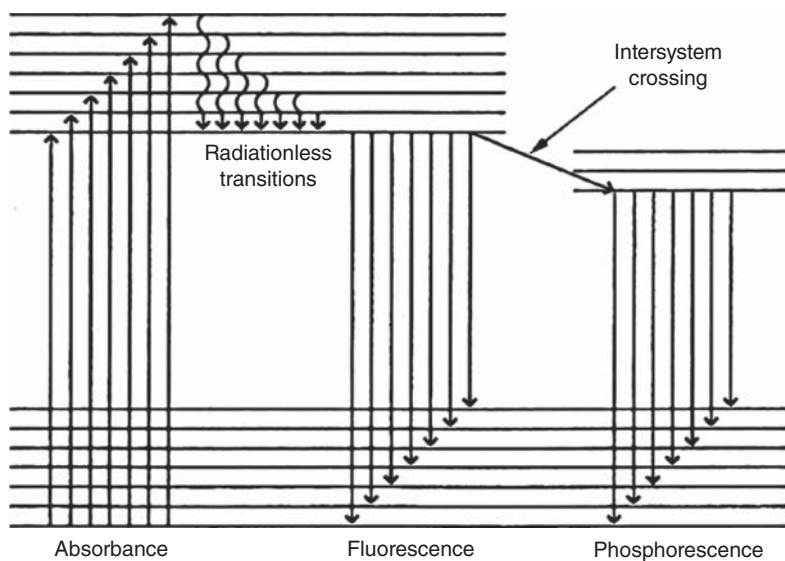


FIGURE 10-11. Schematic energy level diagram for absorption, fluorescence, and phosphorescence. This is often referred to as a Jablonski diagram. Electrons are excited from a ground-state singlet energy level, S_0 , to various vibrational states of a higher electronic energy level, S_1 . All electrons are paired in singlet states, that is, their spins are in opposite directions. Radiationless decay occurs to the ground vibrational energy level of S_1 . As the electrons return to the ground electronic state, fluorescence occurs. In some cases, the excited electrons can move to a triplet state, T_1 , in a radiationless transition (intersystem crossing). Triplet states have two unpaired electrons. The electrons in the triplet state can then return to the ground electronic state with phosphorescence occurring.

the lowest vibrational energy level of the second electronic energy level very rapidly, in 10^{-13} s or less. This generally occurs through a radiationless transition, that is, through the production of heat. Electrons in the second energy level can then decay to the electronic ground state, again in various vibrational energy levels. The emission produced is fluorescence. The decay to the ground electronic state can also occur without the emission of light, that is, by producing heat. Another type of emission is phosphorescence, which will be discussed later.

From Figure 10-11, it is clear that the emitted light must be at a longer wavelength than the absorbed light because the energy change associated with emission can never be greater than the energy change associated with absorption. From this diagram, it is also clear that the number of photons emitted can never exceed the number of photons absorbed. The efficiency of fluorescence is characterized by the *quantum yield*, Q , which is the fraction of photons absorbed that are eventually emitted:

$$Q = \frac{\text{number of photons emitted}}{\text{number of photons absorbed}} \quad (10-7)$$

In practice, the quantum yield can be quite close to 1, but for most molecules that fluoresce well, it is usually in the range of 0.3–0.7. Not all molecules that absorb light produce measurable fluorescence; in fact, most do not. Instead, decay to the ground state is radiationless, via heat production, and the quantum yield is essentially zero. The only common amino acid that has a useful fluorescence spectrum is tryptophan, and none of the common nucleic acids display significant fluorescence.

The measurement of fluorescence (and phosphorescence) is more complex than the measurement of absorption because the sample must be excited at a specific wavelength and the fluorescence observed at a different wavelength. Typically, the emission is detected perpendicular to the excitation beam. The emission spectrum is obtained by keeping the excitation wavelength constant and observing the emission over a range of wavelengths. The fluorescence of tryptophan, an amino acid found in many proteins, is shown in Figure 10-12. Also shown is the absorption spectrum. The emission spectrum in the figure was determined by exciting tryptophan with light at the absorption maximum of 275 nm. The excitation spectrum is obtained by observing the emission at a given wavelength and varying the excitation wavelength. The quantum yield is different for each pair of excitation and emission wavelengths. An absolute measurement of fluorescence would require knowledge of the number of photons absorbed at the excitation wavelength and the number of photons emitted at the emission wavelength. Such measurements are very difficult to make so that in practice the fluorescence intensity is usually expressed in relative terms or in comparison with a standard solution where the quantum yields are well known.

A more quantitative description of fluorescence is useful for understanding the experimental measurements. The number of excited molecules is proportional to the amount of light absorbed, which can be determined from the Beer–Lambert law:

$$I_t = I_0 e^{-2.3 \epsilon(\lambda)cl} \quad (10-8)$$

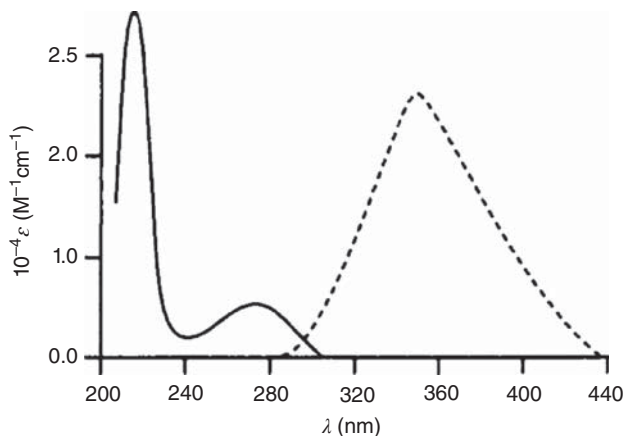


FIGURE 10-12. Absorption and fluorescence spectra of tryptophan. The solid line is the extinction coefficient, and the dashed line is the fluorescence emission in arbitrary units with excitation at approximately 275 nm. D. Freifelder, *Physical Biochemistry*, 2nd edition, Freeman, New York, 1982, p. 539. © 1976, 1982 by Freeman. Used with permission.

Here, I_t is the intensity of the light of incident intensity I_0 after it has passed a distance l through a solution of concentration c with an extinction coefficient of $\epsilon(\lambda)$ at wavelength λ . Fluorescence measurements are made with solutions of low absorbance so that the exponential can be expanded to give

$$I_t = I_0[1 - 2.3\epsilon(\lambda)cl] \quad (10-9)$$

Thus, the concentration of excited molecules is proportional to the light absorbed,

$$I_0 - I_t = 2.3\epsilon(\lambda)cIl_0 \quad (10-10)$$

The fluorescence emitted, F , is the product of the concentration of excited molecules times the quantum yield times the fraction of emitted photons collected by the instrument,

$$F = (\text{constant})Qc \quad (10-11)$$

From the earlier discussion, it is clear that the constant contains several proportionality relationships that are not easily determined, thus necessitating the determination of relative fluorescent intensities rather than absolute intensities. From a practical standpoint, it is important to note that the fluorescence is proportional to the concentration *if* the absorbance of the solution is sufficiently low to permit the expansion of the exponential. Fortunately, this situation is usually quite easy to obtain experimentally.

The discussion thus far is valid for *singlet* electronic states, that is, states in which all of the electron spins are paired. Many complex molecules have electronic states in which two unpaired electrons exist, a *triplet* electronic state. These are at a higher energy state than the ground singlet state and can be at a lower energy than the first

excited singlet state. This is shown schematically in Figure 10-11. In principle, the molecule can move from the excited singlet state to the triplet state, and then decay can occur to the ground state via light emission or the production of heat. The decay via light emission is termed *phosphorescence*. As with fluorescence, the emitted light must be at a longer wavelength than the absorbed light.

What makes phosphorescence different than fluorescence, and how can they be distinguished? In the case of fluorescence, quantum mechanics tells us that the transitions between energy states are allowed, and they occur very rapidly, typically in nanoseconds. In the case of phosphorescence, quantum mechanics tells us that the interconversion from a single state to a triplet state is not allowed, that is, it will not occur readily. The result is that phosphorescence generally takes place in milliseconds or longer. Although this qualitative distinction between fluorescence and phosphorescence is usually sufficient to distinguish between the two, the proof of what is occurring requires determination of the magnetic properties of the excited state. Because the triplet state has unpaired electron spins, its energy will be altered in a magnetic field, whereas this will not happen for fluorescence. At room temperature, DNA does not phosphoresce, but in frozen solutions, thymidine will phosphoresce. We will not explicitly consider phosphorescence further, but it should be remembered that it is very similar to fluorescence; moreover, many biological systems display phosphorescence, perhaps most notable are plant and animal life in the ocean.

Figure 10-11 is a considerable simplification of the real situation where many different electronic states are possible so that multiple electronic energy levels must be taken into account. In addition, many different vibrational modes are possible, so multiple manifolds of vibrational energy levels are generally present. However, the principles enunciated are equally valid for these more complex situations.

10.9 RecBCD: HELICASE ACTIVITY MONITORED BY FLUORESCENCE

As mentioned earlier, part of the mechanism of DNA synthesis involves unwinding DNA. RecBCD is a protein from *Escherichia coli* that functions as a helicase (cf. Ref. 4). It also has an endo- and exonuclease activity, that is, it hydrolyzes DNA. The helicase activity requires the hydrolysis of MgATP to furnish the Gibbs energy necessary for the unwinding process. Monitoring absorbance is not a useful assay for following the unwinding process because of the small changes in absorbance accompanying unwinding relative to the total absorbance. However, a fluorescence assay was developed by Roman and Kowalczykowski (4) that provides a continuous assay of the helicase activity. This assay makes use of a protein that binds specifically to single-stranded DNA, SSB. The binding process stabilizes the single-stranded DNA. Fortunately, when SSB binds to single-stranded DNA, the fluorescence of SSB due to tryptophan is quenched. If SSB is put into an assay mixture containing double-stranded DNA, RecBCD, and MgATP, the SSB will bind to the single-stranded DNA as it is produced. The result is a decrease in the tryptophan fluorescence as more and more single-stranded DNA is formed. This change in

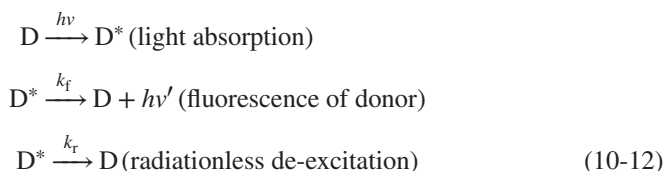
fluorescence is a direct monitor of the rate of DNA unwinding because the binding of SSB to single-stranded DNA is fast relative to the rate of the unwinding reaction. The apparent turnover number for the enzyme determined by this method ranged from 14 to 56 μM base pairs s^{-1} (μM recBCD) $^{-1}$, depending on the experimental conditions.

The mechanism of action of the recBCD helicase is postulated to proceed by the enzyme binding to DNA and unwinding many base pairs before it dissociates from the DNA. An important question to ask is how much unwinding occurs before the helicase dissociates. When an enzyme catalyzes many events during a single binding, the reaction is called *processive*: the greater the amount of unwinding/binding event, the greater the processivity of the enzyme activity. With the assay described earlier, it was possible to quantitate the processivity (5). The processivity increases in a hyperbolic manner as the MgATP concentration increases, reaching a limiting value of 32 kilobases at saturating concentrations of MgATP. The apparent dissociation constant for MgATP deduced from the hyperbolic reaction isotherm is about 40 μM . The processivity was found to be quite sensitive to salt concentration. Thus, a substantial amount of DNA is unwound during each event, ultimately resulting in the synthesis of long stretches of DNA by the polymerase enzyme.

This is but one example of the use of fluorescence to develop an assay for an important biological process, with the subsequent determination of mechanistic aspects of the process.

10.10 FLUORESCENCE ENERGY TRANSFER: A MOLECULAR RULER

Thus far, we have considered only what happens within a single molecule when light is absorbed. However, it is also possible for a molecule in an excited electronic state to transfer the energy above the ground state to another molecule. This phenomenon is termed *Förster energy transfer* (6). This can be readily understood if a simple kinetic scheme is considered. If the energy donor, D, is considered, fluorescence can be described by the following sequence of events:



The k 's are rate constants in this scheme, ν is the light frequency, and * designates the excited electronic state. The “natural” fluorescence lifetime is defined as

$$\tau_0 = \frac{1}{k_f} \quad (10-13)$$

This lifetime cannot be measured experimentally but is a useful theoretical concept.

The fluorescence lifetime is measured by exciting the molecule with short pulses of light and following the decay of fluorescence. According to Eq. (10-12),

$$-\frac{d[D^*]}{dt} = (k_f + k_r)[D^*] = \frac{[D^*]}{\tau_D} \quad (10-14)$$

so that the lifetime determined experimentally is

$$\tau_D = \frac{1}{(k_f + k_r)} \quad (10-15)$$

Eq. (10-14) can be easily integrated with the boundary conditions that $(D^*) = (D_0^*)$ when $t = 0$ and $(D^*) = 0$ when $t = \infty$. The result is

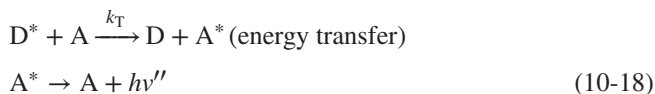
$$[D^*] = [D_0^*] \exp\left(\frac{-t}{\tau_D}\right) \quad (10-16)$$

Thus, a plot of the natural logarithm of the fluorescence intensity versus time is a straight line with a slope of $-(1/\tau_D)$.

The quantum yield for this sequence of events is simply the fraction of molecules fluorescing after excitation, or

$$Q_D = \frac{k_f}{(k_f + k_r)} = \frac{\tau_D}{\tau_0} \quad (10-17)$$

If the frequency of the fluorescence happens to overlap the absorption of a second molecule, the acceptor (A), then some of the energy of de-excitation can be transferred directly to A, which in turn can fluoresce. The scheme in Eq. (10-12) can be expanded to include this possibility:



It should be noted that this energy transfer can occur even if A does not fluoresce, as A could return to its ground state without producing light. The quantum yield of D in the presence of A is now given by

$$Q_{DA} = \frac{k_f}{(k_f + k_r + k_T)} \quad (10-19)$$

The efficiency of energy transfer, E , is

$$E = \frac{k_T}{(k_f + k_r + k_T)} = 1 - \frac{Q_{DA}}{Q_D} = 1 - \frac{\tau_{DA}}{\tau_D} \quad (10-20)$$

where τ_{DA} is the fluorescence lifetime in the presence of the acceptor. In this case,

$$-\frac{d[D^*]}{dt} = (k_f + k_r + k_T)[D^*] \quad (10-21)$$

Thus, the efficiency of energy transfer can be measured quite readily.

This efficiency depends very strongly on the physical separation of the acceptor and donor, and therefore can be used to calculate the distance between A and D. The dependence of the efficiency on the distance between A and D, R , is given by

$$E = \frac{R_0^6}{(R_0^6 + R^6)} \quad (10-22)$$

where

$$R_0(\text{in nm}) = 8.79 \times 10^{-6} (J \kappa^2 n^{-4} Q_D)^{1/6} \quad (10-23)$$

In Eq. (10-23), J is a measure of the spectral overlap of the fluorescence emission of the donor with the absorption of the acceptor, κ is a measure of the relative orientations of the donor and acceptor, and n is the refractive index. Equation (10-22) was first enunciated by Förster, hence the name Förster energy transfer. Because R_0 can be readily calculated, R can be determined directly from measurements of E . Since E depends on the sixth power of the distance, it is a very sensitive measure of the distance. Typical values of R_0 are 3–6 nm so that distances in the range of 1–10 nm are readily accessible to this method. The measurement of energy transfer has proved very useful for providing structural information about biological systems, especially in cases where the molecular structures are not well known.

10.11 APPLICATION OF ENERGY TRANSFER TO BIOLOGICAL SYSTEMS

The use of energy transfer as a molecular ruler in macromolecular systems was elegantly tested by Stryer and Haugland (7). They synthesized a series of proline polymers with a fluorescent donor, dansyl, at one end and an energy acceptor, naphthyl, at the other end. Proline was selected because the polymer is a rigid cylinder whose length can be easily calculated. The number of prolines in the polymer varied from 1 to 12, and R varied from about 1.2 to 4.6 nm. The energy transfer efficiency obeyed Eq. (10-22) exceedingly well, as shown in Figure 10-13, and the experimental values of R agreed well with those determined from molecular models. They also demonstrated that the dependence of R_0 on the degree of overlap of the donor emission with the acceptor absorbance was quantitatively correct.

With the establishment of the methodology in model systems, energy transfer measurements were made in many biological systems. The most critical part of the application is to label a protein or nucleic acid with a fluorescent probe at a specific site, with essentially no labeling at nonspecific sites. This is essential if a reliable

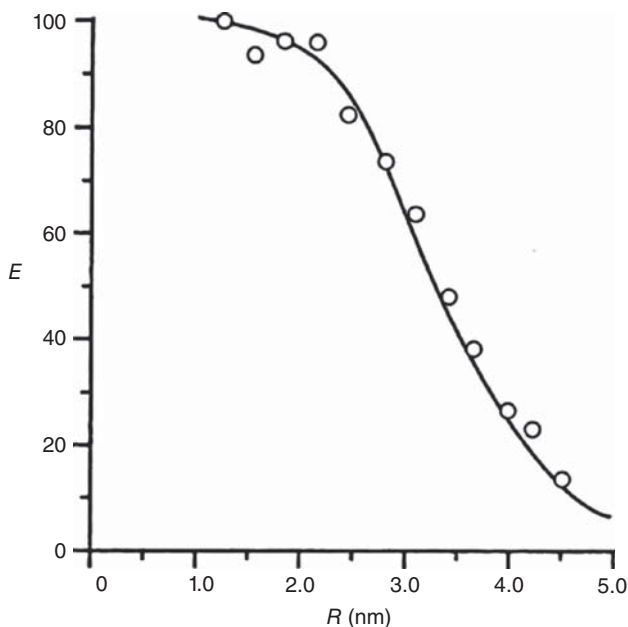


FIGURE 10-13. Percent efficiency of energy transfer, E , in proline polymers of various lengths, R , with a fluorescence donor at one end of the polymer and a fluorescence acceptor at the other end. The solid line corresponds to Eq. (10-22). Reproduced with permission from L. Stryer and R. P. Haugland, *Proc. Natl. Acad. Sci. USA* **58**, 719 (1967).

interpretation of the data is to be made. As a specific example, consider the extensive studies on chloroplast coupling factor 1, CF_1 (8). CF_1 is the soluble part of an enzyme that is critical for ATP synthesis in chloroplasts. Very similar enzymes are in all organisms. In animals, the enzyme is located in the mitochondria. The enzyme is responsible for the synthesis of ATP from ADP and inorganic phosphate, an important physiological reaction. The synthesis occurs when a pH gradient is established across the membrane, and protons are pumped across the membrane as the synthesis proceeds. The coupling of a proton gradient and ATP synthesis is discussed in Chapter 3. Of course, an enzyme must catalyze the chemical reaction in both directions. Thus, the soluble portion of the enzyme catalyzes the reverse reaction, namely the hydrolysis of ATP. It cannot catalyze the synthetic reaction because a proton gradient cannot be created when the enzyme is not on the membrane.

CF_1 is a very complex protein and has five different subunits, with a subunit structure of $\alpha_3\beta_3\gamma\delta\epsilon$. It binds three molecules of ATP or ADP, and the molecule is intrinsically asymmetric. In a series of experiments, fluorescent probes were put in a variety of positions on the molecule, and the distances between the probes were measured. Altogether, more than 30 distances were measured, including the distance of CF_1 sites from the surface of the membrane. These measurements were used to construct a model of the enzymes. The model is shown in Figure 10-14. At the time this work was done, the structure of the enzyme was unknown, and the model summarized

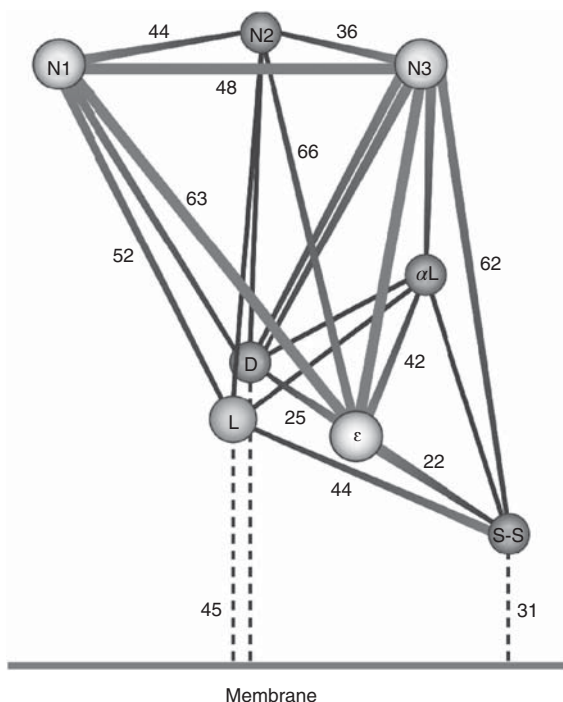


FIGURE 10-14. Spatial arrangement of specific sites on CF₁ determined by fluorescence energy transfer measurements. The sites shown are nucleotide binding sites (N₁, N₂, N₃), a disulfide on the γ polypeptide chain (S-S), specific sulfhydryl groups on the γ polypeptide chain (D, L), a sulfhydryl on the ε polypeptide chain (ε), and an amino group on an α polypeptide chain (αL). A few of the distances, in Å, are shown to establish the scale. The distances to the membrane are maximum values since the lines need not be perpendicular to the membrane surface. Reprinted from R. E. McCarty and G. G. Hammes, *Molecular Architecture of Chloroplast Coupling Factor 1*, *TIBS* **12**, 234 (1987). © 1987, with permission from Elsevier.

some important features of the enzyme, including the distance between nucleotide binding sites and the intrinsically asymmetric nature of the structure. The asymmetric structure is critical for the coordination of the chemical reaction and the pumping of protons across the membrane. Sometime later, the structure of the mitochondrial enzyme was determined by X-ray crystallography (9). The distances determined by energy transfer proved to be consistent with the crystal structure. In fact, the elucidation of structural features by energy transfer measurements is generally useful for complex structures for which the molecular structure has not yet been established.

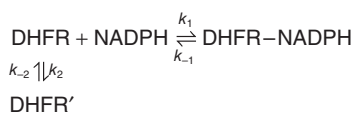
10.12 DIHYDROFOLATE REDUCTASE

As an example of the use of fluorescence to study ligand binding to a protein and enzyme catalysis, we consider the enzyme dihydrofolate reductase (DHFR).

DHFR catalyzes the reduction of 7,8-dihydrofolate by NADPH to give 5,6,7,8-tetrahydrofolate and NADP^+ . Tetrahydrofolate is essential for the biosynthesis of purines, as well as several other important biological molecules. Because of its central position in metabolism, it has been extensively studied. It is also a target for anticancer drugs such as methotrexate. The structures of the enzyme and enzyme–substrate/inhibitor complexes are known, and many mechanistic studies of the enzyme have been carried out (cf. Refs. 10–12). This discussion will be limited to a few facets of the very extensive literature available for this enzyme.

NADPH is fluorescent, with an absorption maximum at about 340 nm and an emission maximum at about 450 nm. In addition, the enzyme contains tryptophan so that if the enzyme is excited at 290 nm, fluorescence is observed with a maximum intensity at 340–350 nm. If NADPH is added to the enzyme, this fluorescence is quenched about 65%, indicating the tryptophan environment is significantly altered when NADPH binds (13). Moreover, because NADPH has a maximum absorbance around 340 nm, significant energy transfer occurs from the excited tryptophan to bound NADPH. This results in a new emission maximum at 420–450 nm. The extent of NADPH binding to DHFR can be quantitatively assessed by measuring the quenching of tryptophan fluorescence. This can also be done by following the fluorescence at 450 nm. Quantitative analysis of the data permits the calculation of the equilibrium dissociation constant, which is 1.1 μM at pH 7.0.

The kinetics of ligand binding can also be measured by mixing NADPH and DHFR rapidly in a stopped-flow apparatus and following the quenching of tryptophan fluorescence or the emission at 450 nm as a function of time. The time course of the reaction indicates a very rapid quenching, occurring in less than a second, followed by a relatively slower reaction, occurring in seconds. The first reaction is the binding of NADPH to DHFR, whereas the second reaction is the interconversion of the free enzyme from a form that does not bind NADPH to a form that does. The overall reaction can be represented by the following mechanism:



A quantitative analysis of the data at pH 7.0 gives $k_1 = 1.7 \times 10^6 \text{ M}^{-1} \text{ s}^{-1}$, $k_{-1} = 2.4 \text{ s}^{-1}$. The equilibrium constant for the transition from DHFR to DHFR' was found to be about 1, with rate constants of approximately $2.5 \times 10^{-2} \text{ s}^{-1}$ (k_2 and k_{-2} ; 11). Note that for this mechanism, the overall equilibrium dissociation constant is

$$\begin{aligned} & \frac{([\text{DHFR}] + [\text{DHFR}'])[\text{NADPH}]}{[\text{DHFR-NADPH}]} \\ &= \left(\frac{k_{-1}}{k_1} \right) (1 + k_2/k_{-2}) = 2.8 \mu\text{M} \end{aligned}$$

The dissociation constant determined from kinetics is only in fair agreement with the dissociation constant determined from direct binding studies.

The overall time course of the enzymatic reaction can be easily monitored by following the absorbance change at 340 nm because NADPH has an absorption peak at this wavelength, whereas NADP⁺ does not. In addition, the individual steps in the catalytic process can be characterized by stopped-flow studies that monitor alterations in the tryptophan fluorescence as the reaction proceeds. A detailed description of these experiments is beyond the scope of this discussion. However, it was possible to determine which substrate binds first to the enzyme, the rate constants for the binding steps, the rate constants for conformational changes within the enzyme–substrate complex, and the rate constants for hydride transfer. This permitted the development of a detailed mechanism for the reaction (11,12).

Another unique use of fluorescence is fluorescence microscopy. With this technique, individual molecules are labeled with fluorescent tags and can be observed with a light microscope (cf. Chapter 6 and Refs. 14, 15). Of course, it is not the single molecule that is seen: it is the light from the molecule. This is analogous to seeing the light from a star, rather than the star itself. With this technique, kinetic and equilibrium properties of individual molecules can be studied. For example, the kinetics of binding of substrates and inhibitors to DHFR and catalysis have been investigated by single-molecule fluorescence microscopy (16).

REFERENCES

1. K. Rosenheck and P. Doty, *Proc. Natl. Acad. Sci. USA* 47, 1775 (1961).
2. R. E. Dickerson and I. Geis, *Hemoglobin: Structure, Function, Evolution and Pathology*, Benjamin/Cummings, Redwood City, CA, 1983.
3. D. G. Anderson, G. G. Hammes, and F. G. Walz, Jr., *Biochemistry* 7, 1637 (1968).
4. L. J. Roman and S. C. Kowalczykowski, *Biochemistry* 28, 2863 (1989).
5. L. J. Roman, A. K. Eggleston, and S. C. Kowalczykowski, *J. Biol. Chem.* 267, 4207 (1992).
6. T. Förster, *Discuss. Faraday Soc.* 27, 7 (1959).
7. L. Stryer and R. P. Haugland, *Proc. Natl. Acad. Sci. USA* 58, 719 (1967).
8. R. E. McCarty and G. G. Hammes, *TIBS* 12, 234 (1987).
9. J. Abrahams, A. Leslie, R. Lutter, and J. Walker, *Nature* 370, 621 (1994).
10. M. R. Sawaya and J. Kraut, *Biochemistry* 36, 586 (1997).
11. C. A. Fierke, K. A. Johnson, and S. J. Benkovic, *Biochemistry* 26, 4085 (1987).
12. G. G. Hammes, S. J. Benkovic, and S. Hammes-Schiffer, *Biochemistry* 50, 10422 (2011).
13. P. J. Cayley, S. M. J. Dunn, and R. W. King, *Biochemistry* 20, 874 (1981).
14. S. Weiss, *Science* 283, 1676 (1999).
15. R. D. Smiley and G. G. Hammes, *Chem. Rev.* 45, 3080 (2006).
16. N. M. Antikainen, R. D. Smiley, S. J. Benkovic and G. G. Hammes, *Biochemistry* 45, 7990 (2006).

PROBLEMS

- 10-1.** Colored pH indicators are dyes that have different spectra for different ionization states. Assume the ionization of a pH indicator with $pK = 6.00$ can be written as



The measured absorbance in a 1-cm cell at an indicator concentration of $2 \times 10^{-5} \text{ M}$ is given in the following table.

λ (nm)	A at pH 3.00	A at pH 9.00
400	0.200	0.000
420	0.300	0.030
440	0.150	0.150
460	0.000	0.200
480	0.000	0.150

For the same concentration of indicator, an absorbance of 0.100 is measured at $\lambda = 400 \text{ nm}$.

- What is the pH of the solution?
 - What is the absorbance at $\lambda = 440 \text{ nm}$?
 - At pH 7.00, for an indicator concentration of $3.00 \times 10^{-5} \text{ M}$ and a 1-cm cell, what is the absorbance at $\lambda = 400, 440$, and 480 nm ?
- 10-2.** A sample of RNA is hydrolyzed and separated into three fractions by column chromatography. Two of the three fractions are pure nucleotides, but the third contains both adenylic and guanylic acids. At pH 7.0, the absorbance of the mixture is 0.305 at 280 nm and 0.655 at 250 nm in 1-cm cells. The molar extinction coefficients for each pure component at pH 7.0 are

	$\epsilon_{280} (\text{M}^{-1} \text{ cm}^{-1})$	$\epsilon_{250} (\text{M}^{-1} \text{ cm}^{-1})$
Adenylic acid	2300	12,300
Guanylic acid	9300	15,700

Calculate the mole ratio of adenine to guanine in the RNA.

- 10-3.** The enzyme alcohol dehydrogenase catalyzes the oxidation of alcohol by NAD^+ to give acetaldehyde and NADH. NADH has an absorption maximum at 340 nm with an extinction coefficient of $6.20 \times 10^3 \text{ M}^{-1} \text{ cm}^{-1}$, whereas NAD^+ and the other reactants do not absorb significantly at 340 nm. Consequently, the oxidation reaction can be conveniently monitored by following the increase in absorbance at 340 nm.

- a. What is the rate of production of NADH if it is observed that the rate of absorbance increase in a 1-cm path length cell is 0.05/min?
- b. When an excess of enzyme is added to NADH, the absorbance at 340 nm decreases by 13%. What is the extinction coefficient of NADH at 340 nm when it is bound to the enzyme?
- c. The difference in extinction coefficient can be used to determine the binding constant for the association of NADH and the enzyme. The following data were obtained for the difference absorbance when NADH is added to 20 μM enzyme in a 1-cm path length cell.

(NADH) _{total} (μM)	ΔA (340 nm)
8.33	0.004
13.50	0.006
20.0	0.008
45.0	0.012

Calculate the binding constant from these data.

- 10-4.** When complementary strands of deoxyoligonucleotides are mixed, they form a double-stranded DNA at low temperatures but dissociate to single strands as the temperature is raised. This can be observed by a decrease in the absorbance of the solutions at 260 nm. The equilibrium for this process can be written as



with an equilibrium constant $K = (\text{D})/[(\text{S}_1)(\text{S}_2)]$. Data for two experiments are given below. In one case, equal length strands and concentrations of oligoA and oligoT were mixed, whereas, in the other case, the same length strands and concentrations of oligoG and oligoC were mixed.

- a. Assume that the molecules are duplexes at the lowest temperatures for which data are given and single strands at the highest temperatures. Calculate the fraction present as duplex at each temperature. Determine the melting temperature of each duplex, that is, the temperature at which half of the duplex has been converted to single strands.
- b. What do these results tell you about the relative stability of A–T and G–C pairs?
- c. Estimate the melting temperature of complementary oligonucleotides that are 50% A–T and 50% G–C, with the same length as in the aforementioned experiments. (Your result is not exact, but methods are available for calculating the melting temperatures of short DNA; cf. Chapter 3.)

$T(^{\circ}\text{C})$	A (oligoA–T)	$T(^{\circ}\text{C})$	A (oligoG–C)
5	0.711	50	0.780
10	0.720	55	0.785
15	0.732	60	0.812
20	0.740	65	0.836
25	0.767	70	0.862
30	0.801	75	0.880
35	0.846	80	0.896
40	0.874	85	0.924
45	0.891	90	0.948
50	0.903	95	0.975
55	1.003		

- 10-5.** Pyrenylmaleimide is a convenient fluorescent probe for labeling sulfhydryl residues on proteins. The fluorescence lifetime of pyrene on a protein sulfhydryl was measured by determining the relative fluorescence intensity after excitation of pyrene with a flash lamp. Typical data are given below.

Relative fluorescence	Time (ns)
0.716	20
0.513	40
0.367	60
0.264	80
0.189	100

- a. Determine the fluorescence lifetime.
 - b. If the quantum yield is 0.7, what is the natural lifetime?
 - c. When a ligand is bound to the protein, the fluorescence lifetime is 10% shorter. What might be the cause of this change?
- 10-6.** Bovine rhodopsin is a photoreceptor protein that is an integral part of the disc membranes of retinal rod cells and plays a key role in vision. It has a tightly bound 11-*cis*-retinal that has a strong absorbance at about 500 nm. As part of the vision cycle, the retinal is bleached by conversion to the *trans*-isomer. The protein has a molecular weight of 28,000–40,000. Three sites were labeled on the protein with fluorescent probes, sites A, B, and C. Fluorescence resonance energy transfer was measured between these three sites and the retinal and between the three sites themselves [C.-W. Wu and L. Stryer, *Proc. Natl.*

Acad. Sci. USA **69**, 1104 (1972)]. Some of the results are summarized in the following table.

Energy donor	Energy acceptor	Transfer efficiency	$R_0(\text{\AA})$
A	11- <i>cis</i> -Retinal	0.09	51
B	11- <i>cis</i> -Retinal	0.36	52
C	11- <i>cis</i> -Retinal	0.12	33
A	B	0.90	51
A	C	0.92	48
B	C	0.92	47

- a. Calculate the distances between these six sites.
- b. A protein of molecular weight 28,00–40,000 that is spherical has a radius of 40–45 $\text{\AA}$. What can you say about the shape of rhodopsin? Sketch a model for the molecule based on the distances that have been measured.

Circular Dichroism, Optical Rotary Dispersion, and Fluorescence Polarization

11.1 INTRODUCTION

Most biological molecules possess molecular asymmetry, that is, their mirror images are not identical. Such molecules are said to be *chiral*. Probably the most well-known example is a carbon atom that is tetrahedrally bonded to four different atoms or groups of atoms. This example is described in organic chemistry textbooks and will not be dwelled on here. Less obvious examples of importance in biology are macromolecules that possess chirality. For example, helices can be wound in a left- or right-hand sense. The most common helix in proteins, the α -helix, is wound in a right-hand sense. Although most polynucleotides are wound in a right-hand sense, helices that wind in a left-hand sense also exist.

Chiral molecules can be distinguished by their interactions with polarized light. As we have discussed in Chapter 7, a light wave can be described as an electromagnetic sine wave with a characteristic frequency. In this description, a wave generating an electric field is perpendicular to a wave generating a magnetic field. If light is *unpolarized*, these waves are oriented randomly in space, as shown schematically in Figure 11-1. Looking along the electromagnetic wave, the random orientations look like the spokes of a wheel. For *linearly*, or *plane*, *polarized* light, the wave is oriented in only one direction, as depicted in Figure 11-1. Although the polarized light in the figure looks like a line (linearly polarized), the electric field wave is projected as a plane (plane polarized) so that the two terms are used interchangeably. When a chiral molecule interacts with plane polarized light, the plane is rotated, with the direction and amount of rotation depending on the characteristics of the molecule.

Circularly polarized light sweeps out a circle as the electric wave propagates. The circle can be either right handed or left handed, as shown in Figure 11-1. The convention used is that if the circle moves clockwise, it is right handed, whereas if it moves counterclockwise, it is left handed. The interaction of circularly polarized light with chiral molecules will alter the circularly polarized light. Elliptically polarized light is also used and is essentially the same as circularly polarized light, but an ellipse is swept out by the electric field wave instead of a circle. We will not dwell on how polarized light is produced, except to note that light can be polarized when it is passed

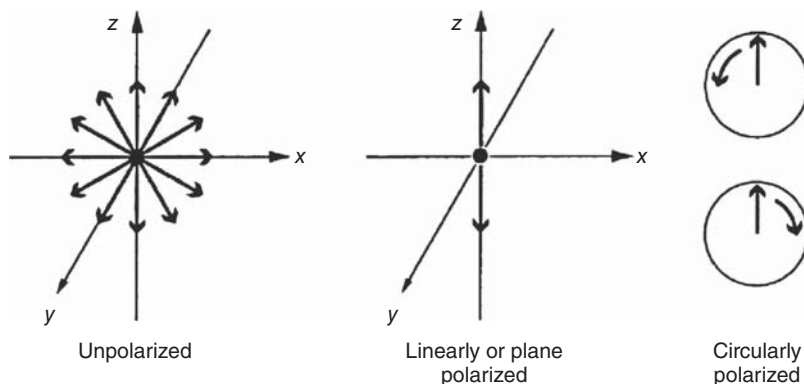


FIGURE 11-1. Diagrams of the electric field components of unpolarized, linearly or plane polarized, and circularly polarized light. The light is moving along the y-axis. The arrows indicate the directions of the electric field. For unpolarized light, all directions occur, whereas for linearly or plane-polarized light only the z direction is found. For circularly polarized light, the direction of rotation can be clockwise or counterclockwise.

through certain materials. A pair of polaroid sunglasses is a good example of how light can be polarized.

The interaction of a chiral molecule with polarized light is very specific and has proved to be an important method for characterizing both small molecule and macromolecular structures. Small molecule examples should be already well known to you. Amino acids in most biological systems are *levo* rotary (L), that is, they rotate plane polarized light to the left, and sugars have various optical isomers. Enzymes, in fact, can distinguish between optical isomers and typically will only react with a single or a restricted group of small molecule isomers. Essentially, two types of measurements are commonly made to determine the effects of molecules on polarized light: optical rotation and circular dichroism (CD). Optical rotation is a measure of the rotation of linearly polarized light by a molecule, and the wavelength dependence of the optical rotation is called optical rotary dispersion (ORD). CD, on the other hand, is the difference in absorption of left-hand and right-hand circularly polarized light. These effects are relatively small but can be measured readily with modern instrumentation. For a typical protein or nucleic acid with a 100 μM solution of chromophore, the polarized light plane with a wavelength around the electronic absorption maxima is rotated about 0.001–0.1 degrees for a sample 1 cm thick. For CD, the difference in absorption coefficients is about 0.03–0.3% of the total absorption.

11.2 OPTICAL ROTARY DISPERSION

The first studies of the optical properties of biological macromolecules were done with ORD because adequate equipment was not available for CD measurements. At the present time, the opposite is true: CD is the method of choice. Nevertheless, it is instructive to discuss both methods. As we shall see later, they are intimately related.

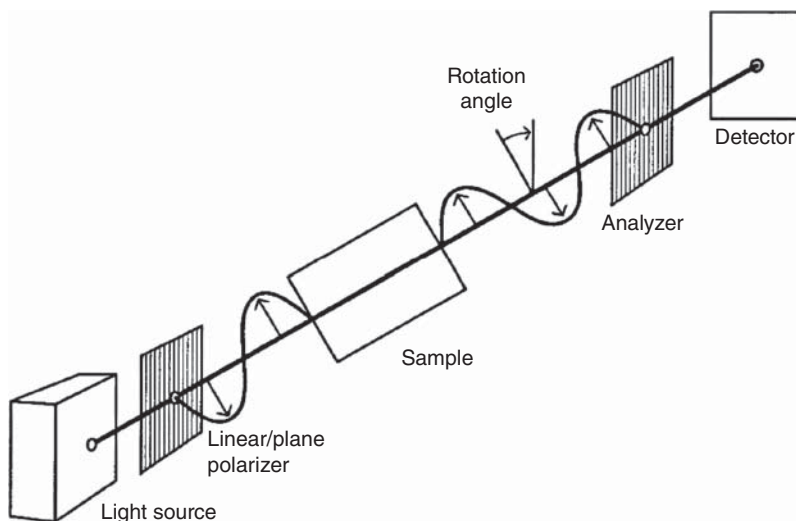


FIGURE 11-2. Schematic diagram of an ORD apparatus. The light is linearly polarized and the polarization plane is rotated by the sample. A second polarizer, the analyzer, can be rotated to obtain the maximum (or minimum) light intensity. The angle of rotation of the analyzer is a direct measure of the angle of rotation by the sample. For CD, circularly polarized light is used, and the difference in absorbance between left and right circularly polarized light is measured.

A typical experimental setup for measuring ORD is shown in Figure 11-2. The principle is that plane-polarized light is passed through the sample, and the rotation of the plane of polarization of the light by the sample is measured. The analyzer for the emergent beam is an element that polarizes the beam. It can be rotated until the light intensity after the analyzer disappears. Its polarization axis is then perpendicular to the direction of polarization of the beam emerging from the sample. The amount of rotation required for this disappearance is a direct measure of how much the original beam was rotated by the sample. Clockwise rotation is assigned as positive rotation and counterclockwise as negative rotation. The physical origin of this rotation is the fact that the sample is *circularly birefringent*, that is, the refractive index of the sample is different for left-hand circularly polarized light and right-hand circularly polarized light. A plane-polarized beam of light can also be described as two opposite circularly polarized beams. A difference in the refractive index means that the speed at which the light passes through the sample is different for right-hand and left-hand circularly polarized light. This means the two polarized beams get out of phase as they pass through the sample. This phenomenon is described in Chapter 7. This phase difference is manifested as a rotation of the light.

If the measured angle of rotation is α , then the specific rotation, $[\alpha]$, is defined as

$$[\alpha] = \frac{\alpha}{dc} \quad (11-1)$$

where c is the concentration in g/cc and d is the path length in decimeters. The origin of this definition lies in the history of optical rotation: the first measurements of optical rotation were made in cells that were 10 cm (1 dm) long. Instead of the specific rotation, data are often reported as the molar rotation, $[\phi]$, which is defined as

$$[\phi] = \frac{100\alpha}{lM} \quad (11-2)$$

where l is the path length in centimeters and M is the concentration in mol/l. (For macromolecules, the concentration is often given in terms of the monomers.)

11.3 CIRCULAR DICHROISM

Materials display CD because the absorbance of left and right circularly polarized light is different for molecules with molecular asymmetry. One measure of the CD is simply the difference in absorbance of the material for left and right circularly polarized light, ΔA :

$$\Delta A = A_L - A_R \quad (11-3)$$

Circular dichroism can arise only in the spectral region where absorbance occurs—if the absorbance is essentially zero, there cannot be a measurable difference. For ORD, on the other hand, large rotations are measured in the regions where the sample absorbs light, but measurable rotation also occurs when the absorbance is essentially zero. This is helpful, for example, when looking at the optical activity of substances such as sugars, for which the absorbance is in the far ultraviolet. The large optical rotations in regions of high absorption are called Cotton effects.

As previously discussed, optical rotation arises because of the phase shift of the circularly polarized light that leads to circular birefringence. Circular dichroism arises because not only does the phase shift but there also is a differential decrease in the amplitude for right and left circularly polarized light. This leads to elliptical polarization. As with optical rotation, circular dichroism can be either positive or negative. Circular dichroism is reported either as a differential extinction coefficient,

$$\Delta\epsilon = \epsilon_L - \epsilon_R = \frac{\Delta A}{lM} \quad (11-4)$$

or more often as the ellipticity, θ :

$$\theta = 2.303\Delta A \frac{180}{4\pi} \text{ degrees} \quad (11-5)$$

As before, the molar ellipticity is

$$[\theta] = \frac{100\theta}{lM} \quad (11-6)$$

Despite this definition of $[\theta]$, it is usually reported in the units of $\text{deg cm}^2 \text{ dmol}^{-1}$ because of the historical precedents. The differential extinction coefficient and the specific rotation are related by the relationship

$$[\theta] = 3300\Delta\epsilon \quad (11-7)$$

The numerical coefficient follows from the Beer–Lambert law for absorption and the various relationships in previous equations.

Since ORD and CD arise from the same physical phenomena, namely the effect of molecular asymmetry on polarized light, one might imagine that the two should be closely related. In fact, these two measurements are not independent: if the optical rotation is known, the circular dichroism can be calculated and vice versa. Although the formal equations for these transformations are well known, in practice the mathematics are sufficiently difficult so that this transformation is rarely done. In fact, CD is the experimental method of choice for a variety of reasons, mostly because circular dichroism is only found at absorption bands so that the interpretation of the spectra is somewhat easier.

11.4 OPTICAL ROTARY DISPERSION AND CIRCULAR DICHROISM OF PROTEINS

The simplest model for the optical activity of a protein is to assume that the optical activity is the sum of the optical activity of the individual amino acids. This model is clearly incorrect: the sum of the optical activity of the amino acids is usually very small relative to the measured optical activity of a protein. This observation is not surprising: We saw previously that the ultraviolet spectra of proteins were primarily due to the peptide bonds. Similarly, the optical activity of proteins is primarily due to the macromolecular structure itself. In fact, specific protein structures have characteristic ORD and CD spectra.

Three of the fundamental structures of proteins are the α -helix, the β -sheet, and the random coil (cf. Chapters 3 and 9). The α -helix is a right-hand helix stabilized by short-range hydrogen bonds between backbone peptide bonds, whereas the β -sheet structure is composed of parallel polypeptide chains, also stabilized by hydrogen bonds between the backbone peptide bonds (Figs. 3-4 and 3-5). The random coil is envisaged as a random arrangement of the backbone although it is rarely random—irregular might be a better term. The ORD and CD spectra of many different homopolypeptide chains have been determined under experimental conditions where the structures are known. The results are summarized in Figure 11-3 (1). Within relatively small deviations, the spectra are the same for the common structures of the homopolypeptides. Note that, as expected, the CD is centered around the ultraviolet absorption spectra of the proteins. Also note that the spectra are reasonably distinct for these three different structures.

If the assumption is made that only α -helix, β -sheet, and random coil structures are present and these three structures have known optical properties, it should be possible to develop algorithms to deconvolute the ORD and CD spectra of a given protein.

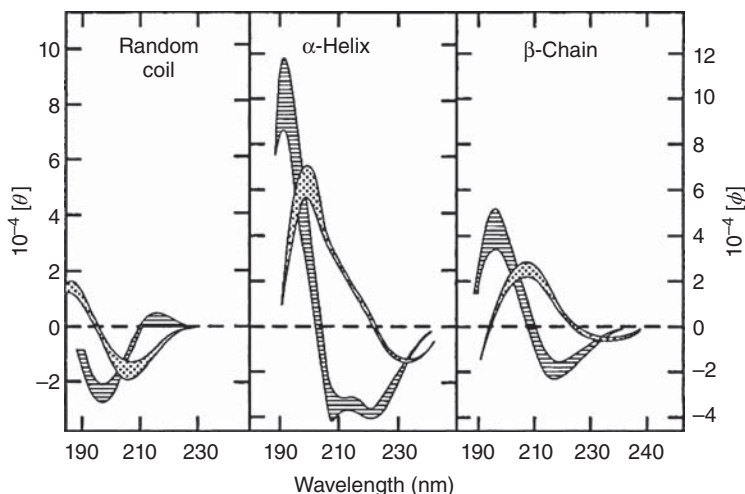


FIGURE 11-3. ORD (dotted) and CD (crosshatched) of homopolypeptides in the random coil, α -helix, and β -sheet conformations. The range of values found is indicated by the shaded areas. Reproduced with permission from W. B. Gratzer and D. A. Cowburn, *Nature* **222**, 426 (1969).

The implicit assumption in these algorithms is that ORD and CD spectra of the three different structures are additive; that is, the presence of one structure does not influence the contributions from another structure. Thus, measurement of the ORD or CD spectrum over the wavelength region of 175–250 nm would permit the calculation of the fraction of the protein structure present in each of the three conformations. A number of attempts have been made to do this, and the results have been compared with known crystal structures (cf. Ref. 2). In some cases, crystal structures have been used to derive a set of self-consistent parameters, assuming that the crystal and solution structures are identical. Reasonable agreement between the calculated structures and the actual structures is often found, but the calculations are not quantitative for several reasons. First, the optical properties of homopolymers are not accurate predictors of the optical properties of the protein. The ORD and CD spectra depend on the microenvironment of each residue, so that they will certainly be dependent on the specific sequence and size of the protein. Second, protein structures are not so simple that their domains can be described by the three relatively simple structures assumed for the algorithms. Other structures that can influence CD and ORD spectra include various helices other than the α -helix, disulfide bonds, and various β -turns. Many of the fitting algorithms include structures other than the three discussed here. Finally, the side chains of amino acids influence the spectra—this is particularly true at wavelengths above about 250 nm where tyrosine, phenylalanine, and tryptophan have strong absorption bands. Prosthetic groups, such as hemes, would have a similar effect at the wavelength where they have strong absorption bands. Nevertheless, ORD and CD spectra are useful semiquantitative predictors of protein structure and

also serve as good indicators of structural changes that may occur as experimental conditions are changed (cf. Ref. 3, 4).

11.5 OPTICAL ROTATION AND CIRCULAR DICHROISM OF NUCLEIC ACIDS

The bases of nucleic acids are not intrinsically optically active; however, the sugars are optically active and can induce optical activity in the bases. This optical activity is relatively small, but as might be anticipated from the discussion of proteins, the ORD and CD spectra of nucleic acids are not just the sum of those of the monomers. In fact, CD and ORD are very good indicators of secondary structure such as helices and are the method of choice for following changes in secondary structure. As an example, consider polyadenylic acid (polyA) (5). As shown in Figure 11-4, at neutral pH and room temperature, native polyA has a very large ellipticity due to the fact that it exists in a helical structure. If the structure is destroyed by raising the temperature or going to extremes of pH, the ellipticity essentially disappears. Also shown in the figure is the ellipticity of adenylic acid, which is quite small. The ORD and CD of

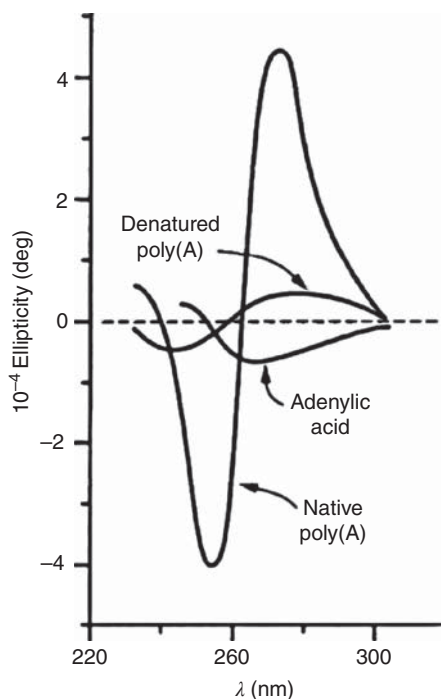


FIGURE 11-4. Circular dichroism of adenylic acid, denatured polyadenylic acid, and native polyadenylic acid. D. Freifelder, *Physical Biochemistry*, 2nd edition, Freeman, New York, 1982, p. 594. © 1976, 1982 by Freeman. Used with permission.

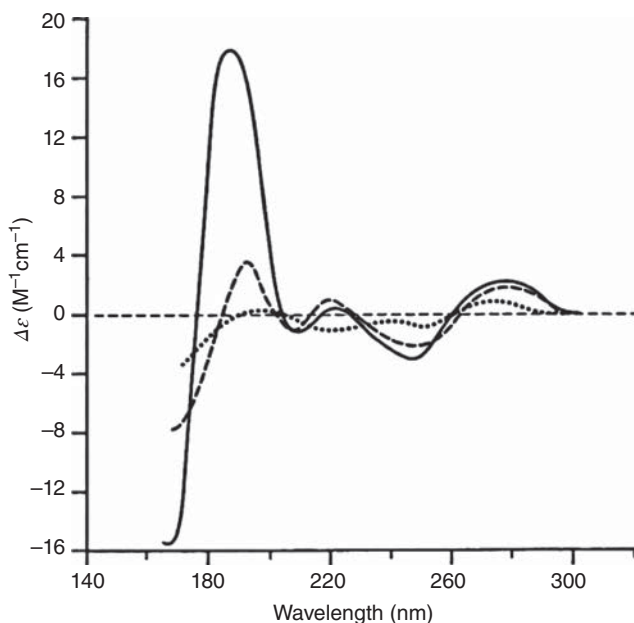


FIGURE 11-5. The CD of *E. coli* DNA in its native form at 20°C (—), heat denatured form at 60°C (---), and the average CD of the four deoxynucleotides (...). From C. A. Sprecher and W. C. Johnson, Jr., *Biopolymers* **16**, 2243 (1977). Reprinted with permission of Wiley © 1977.

single-stranded polynucleotides have been extensively studied. The large ellipticity is primarily due to the interactions of the nearest neighbors, that is, adjacent bases. They are stacked on top of each other, and this stacking is primarily responsible for the altered optical properties of the helical structures. To a good approximation, the ORD and CD properties of a base are influenced only by their nearest neighbors. Consequently, the spectra of polymers can be readily simulated by adding together the spectrum of each base, taking into account nearest neighbor interactions only.

The situation is much more complex for double-stranded nucleic acids such as DNA, and structural predictions are much more difficult, but CD and ORD are still useful tools. As an example, the CD of *E. coli* DNA is shown in Figure 11-5 for both the native and denatured forms (6). Although structural predictions from ORD and CD spectra are difficult, it is true, nevertheless, that specific structures of DNA and RNA have very characteristic spectra, and changes in the spectra are good monitors of structural changes (cf. Ref. 3, 4).

11.6 SMALL MOLECULE BINDING TO DNA

DNA is an obvious target for drug intervention since disruption of its structure clearly will have significant biological implications. The negatively charged phosphate

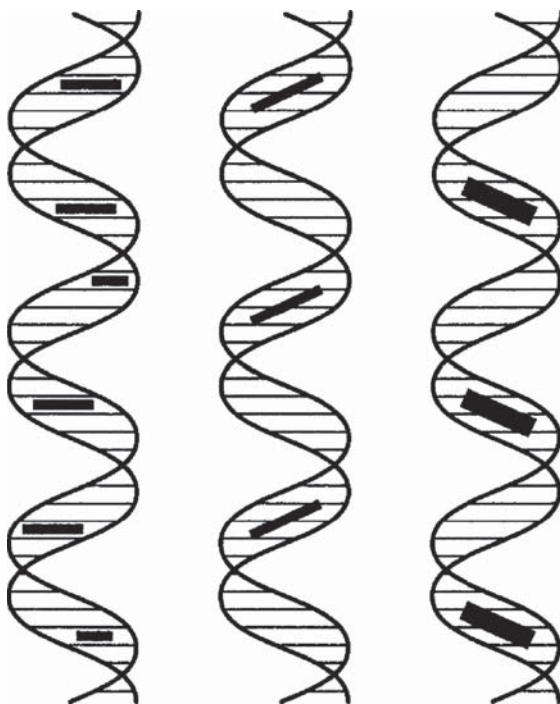


FIGURE 11-6. Schematic diagram of the intercalation of ligands between the bases, and in the minor and major grooves of a DNA double helix. From M. Ardhammar, B. Norden, and T. Kurucsev in *Circular Dichroism: Principles and Applications* (N. Berova, K. Nakanishi, and R. W. Woody, eds.), 2nd edition, Wiley, New York, 2000, p. 746. Reprinted with permission of Wiley © 2000.

groups along the backbone structure of DNA can bind small molecules through electrostatic interactions, but the major sites of drug binding in double-stranded DNA are the major and minor grooves in the double helical structure and intercalation between the bases. These three modes of binding are shown schematically in Figure 11-6 (7). Many of the drugs used contain aromatic rings, and as a specific example of such binding, we consider the interaction of DNA with acridine orange, even though acridine orange is not actually a drug (8).

The structure of acridine orange is shown in Figure 11-7. It is not optically active by itself, but binding to DNA induces optical activity into the molecule. This is because binding causes the electronic energy states of achiral acridine orange to be coupled with the electronic energy states of chiral DNA. Induced optical activity is quite common when achiral small molecules bind to chiral macromolecules. The CD spectra of acridine orange bound to DNA are shown in Figure 11-7 for various ratios of [bound dye]/[DNA] (8). As expected, the optical activity is centered around the electronic absorption band of acridine orange, which has a maximum at about 430 nm. At very low values of [bound dye]/[DNA], a negative value of $\Delta\epsilon$

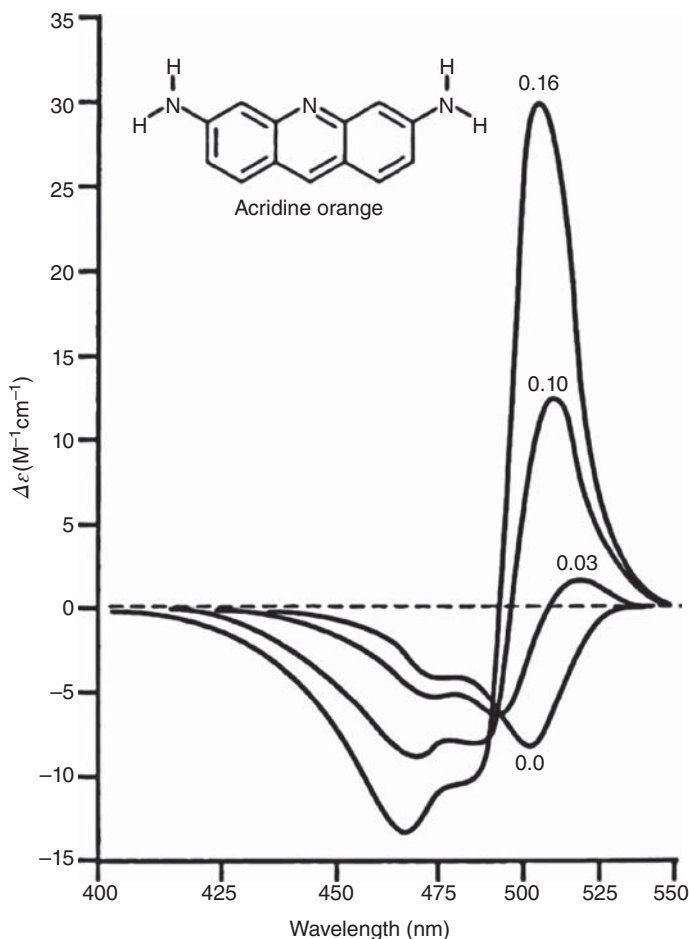


FIGURE 11-7. Circular dichroism per mole of acridine orange bound to DNA at the indicated values of $[\text{bound dye}]/[\text{DNA}]$. The total dye concentration is 10^{-5} M. Reprinted in part with permission from D. Fornasiero and T. Kurnsev, *J. Phys. Chem.* **85**, 613 (1981). © 1981 by American Chemical Society.

is observed, but as this ratio increases, a large positive band develops. At the highest ratio shown, both strong negative and positive bands are observed. A molecular interpretation of these results has been developed. At very low concentrations of acridine orange, the binding is through intercalation between the bases, with the acridine orange ring structure parallel to the bases. This produces a negative induced circular dichroism with a maximum $\Delta\epsilon$ of about $-8\text{ M}^{-1}\text{ cm}^{-1}$. As the concentration of the ligand increases, the dye binds to the groove which induces a positive CD with a maximum value of about $60\text{ M}^{-1}\text{ cm}^{-1}$. These two induced CD signals are enhanced at higher dye concentrations due to interactions between the electronic energy levels of the bound dye molecules.

This example illustrates the exquisite sensitivity of the CD spectrum to the nature of the binding process and to the secondary structure of the DNA. Studies with many different types of ligands have been carried out (cf. Ref. 7).

11.7 PROTEIN FOLDING

Understanding protein structure in molecular terms is a long-standing goal of biochemistry. While great progress has been made toward this goal, we still cannot predict the three-dimensional structure of a protein from a knowledge of the one-dimensional sequence of amino acids that make up a protein. One approach to this end is to study the folding and unfolding of proteins with the rationale that understanding these processes will provide a better understanding of protein structure. In addition, understanding the folding of proteins is of great physiological significance because proteins are synthesized unfolded *in vivo* and must be folded into specific structures in order to perform their biological activities. Protein folding already has been briefly discussed in Chapter 3. Many reviews of protein folding are available (cf. Ref. 9, 10), and only one particular example is presented here. CD and ORD are particularly useful tools for the study of protein folding and unfolding because they are extremely sensitive to the secondary and tertiary structure of proteins so that large changes are generally seen as a result of the folding process.

Traditionally, proteins were thought to be synthesized and then immediately folded into their biologically active form. However, in recent years, a number of examples have emerged in which the unfolded form exists *in vivo* until it is folded into its biologically active form by specific conditions in the external environment such as the presence of a specific ligand. The biological significance is not understood, but the evidence for such structures is quite convincing. An example of this phenomenon is the protein associated with the ribozyme ribonuclease P (11). Ribonuclease P is a ribonucleic acid (RNA) catalyst, previously described in Chapter 6. Its activity and/or specificity is enhanced by the binding of the catalytic RNA to a protein, and the protein is essential for *in vivo* catalysis (12). We will not discuss the catalytic reaction itself which is the processing of the 5'-leader sequences from precursor tRNA. This ribozyme is quite ubiquitous in nature: in bacteria, the complete catalytic unit consists of a single RNA of about 400 nucleotides and a single protein of about 120 amino acids.

When the protein from *B. subtilis* is isolated, it is unfolded at physiological temperatures in the absence of ligands (13). This can be ascertained from the CD spectrum in sodium cacodylate (pH 7), which does not bind to the protein. The spectrum of the protein in the far and near ultraviolet is shown in Figure 11-8. The far-ultraviolet spectrum is due to the protein backbone whereas the near-ultraviolet spectrum is primarily due to tyrosine residues in the protein. Included in Figure 11-8 are the spectra in the presence of 10 or 20 mM sulfate. An obvious change in the spectrum occurs. The far-ultraviolet spectrum is consistent with the known crystal structure of the protein, which contains α -helices and β -sheets, and the near-ultraviolet spectrum changes because the electronic environments of the tyrosines are altered.

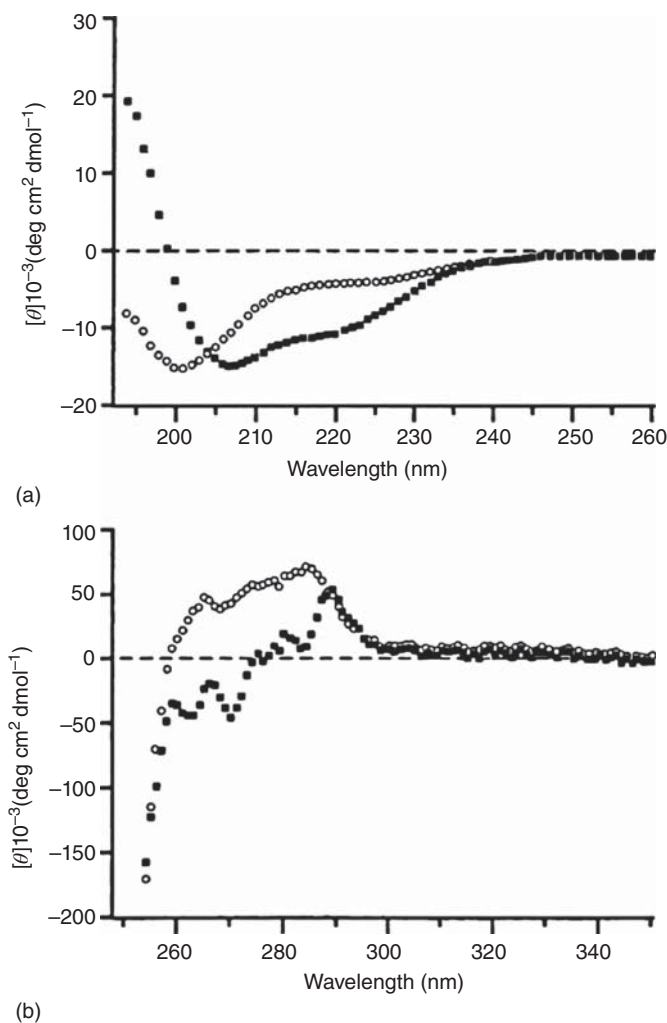


FIGURE 11-8. Circular dichroism of P protein in the presence (○) and absence (■) of 10 (b) or 20 (a) mM sulfate at pH 7.0 in 10 mM sodium cacodylate. Reprinted in part with permission from C. J. Henkels, J. C. Kurz, C. A. Fierke, and T. G. Oas, *Biochemistry* **40**, 2777 (2001). © 2001 by American Chemical Society.

The change in structure can be quantified by measuring the change in the CD at 222 nm as a function of the ligand concentration. This is shown for a variety of ligands in Figure 11-9. A variety of anions cause the protein to go from a disordered structure to a more folded structure, with highly charged anions such as polyphosphates being more effective than monovalent ions such as chloride. The stoichiometry of the ligand binding was determined for pyrophosphate and was found to be 2 pyrophosphates/protein. A simple model that explains the observations is that the ligand binds

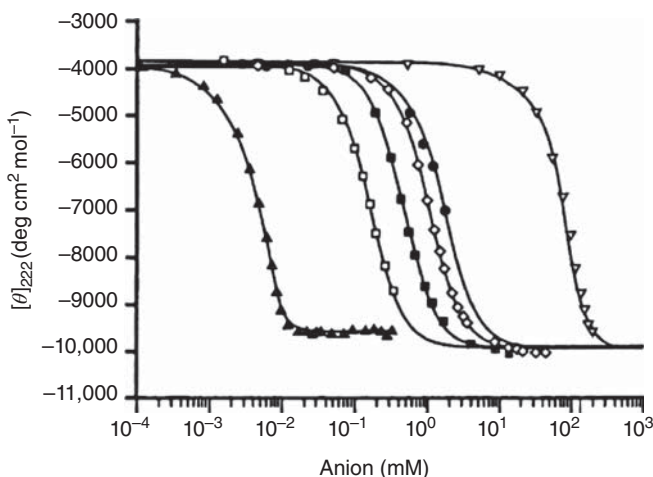
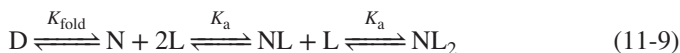


FIGURE 11-9. Anion-induced folding transitions of the P protein followed by the change in the CD signal at 222 nm at 37°C. The anions are dCTP (▲), CMP (□), sulfate (■), phosphate (◇), dCMP (●), and formate (▽). Selected data from C. J. Henkels, J. C. Kurz, C. A. Fierke, and T. G. Oas, *Biochemistry* **40**, 2777 (2001). Figure courtesy of Dr. C. J. Henkels, Duke University. Reproduced with permission.

to the folded protein but not the unfolded protein. The apparent equilibrium constant, K_{app} , for the ratio of the denatured (unfolded) protein, D, to the native state (folded) protein, N, can be obtained directly from the data in Figure 11-9 if the assumption is made that only two forms of the protein exist, folded and unfolded. For example, the apparent equilibrium constant is unity when the change in the CD is half of the total change. The apparent equilibrium constant at any point on the curve is given by

$$K_{\text{app}} = \frac{(\theta_L - \theta_D)}{(\theta_N - \theta_L)} = \frac{\Sigma N}{\Sigma D} \quad (11-8)$$

where θ 's are the absolute values of the circular dichroism in the figure for a given ligand concentration, L, the denatured form, D, and the native form, N. The summation is over all liganded forms of the denatured and native proteins. If the ligand is assumed to bind to two independent and equivalent sites, the folding mechanism can be written as



This mechanism gives

$$\begin{aligned} K_{\text{app}} &= \frac{[N] + [NL] + [NL_2]}{[D]} = \frac{[N](1 + [NL]/[N] + [NL_2]/[N])}{[D]} \\ &= \frac{[N]}{[D]} (1 + 2K_a[L] + K_a^2[L]^2) = K_{\text{fold}}(1 + K_a[L])^2 \end{aligned} \quad (11-10)$$

where $K_{\text{fold}} = [N]/[D]$ and K_a is the microscopic association constant for the binding of ligand to the native protein ($K_a = [NL]/(2[N][L]) = 2[NL_2]/([NL][L])$). The concept of microscopic association constants is considered more extensively in Chapter 17. The two constants K_{fold} and K_a cannot be determined independently from the data in Figure 11-9. However, an independent method, which will not be described here, gives an estimated value of 0.0071 for K_{fold} at 37°C. With this value in hand, it is possible to calculate the microscopic association constants for the ligands.

This example illustrates two important concepts. First, it shows that measurement of the CD provides a means of following the transition between the denatured and native state of proteins. Second, it shows how the binding of a ligand to the native state can convert the denatured protein to its native structure. (Coupled equilibria are common in biological systems.) These concepts are of both practical and theoretical value for studying and understanding biochemical processes.

11.8 INTERACTION OF DNA WITH ZINC FINGER PROTEINS

Zinc finger motifs are found in proteins that regulate transcription. As the name implies, the structure is a finger consisting of a loop of β -structure with two cysteines and an α -helix with two histidines. The four side chains of these amino acids tightly bind a zinc ion, as shown schematically in Figure 11-10. Transcription factors contain multiple “fingers,” and the α -helix of each finger binds in the major groove of double-stranded DNA, thereby regulating transcription. The interaction of zinc finger proteins with DNA has been extensively studied. For example, the interaction

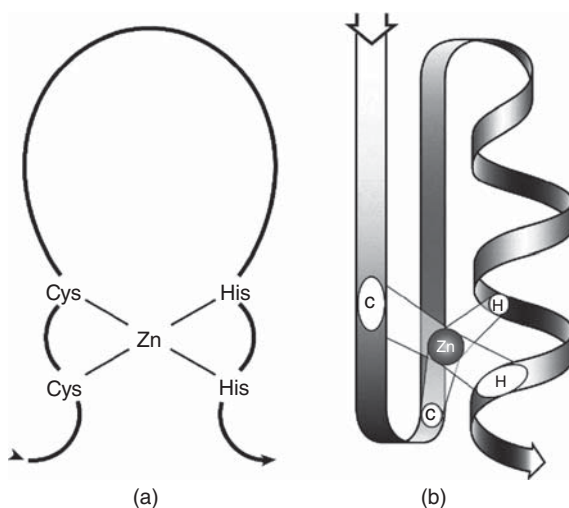


FIGURE 11-10. Schematic drawing of the Zn finger structure. (a) Coordination of Zn to cysteine and histidine ligands. (b) Secondary structure. Reprinted from R. M. Evans and S. M. Hollenberg, Zinc fingers: gift by association, *Cell* **52**, 1(1988). © 1988, with permission of Elsevier.

of a transcription factor that regulates the expression of the protein metallothionein in mouse and human cells with DNA has been studied using CD (14,15). The transcription factor contains six zinc fingers, which were isolated from the rest of the transcription factor. Preparations contained three to six fingers. The six-finger protein binds three zinc ions very tightly, and far-ultraviolet CD spectra showed that the conformation of the protein was unchanged once the three tight binding sites were occupied by zinc, although three additional zinc ions were bound more weakly.

The binding of the isolated zinc fingers to a 23-mer DNA that contains the sequence involved in the regulatory process was studied with CD. As the number of zinc fingers was increased from 3 to 6, a negative ellipticity at about 210 nm decreased in magnitude (approached zero) with a concomitant increase in a positive ellipticity at 280 nm. This change is not due to the protein, whose CD spectrum was subtracted from the experimentally determined spectrum. It is due to a change in the DNA structure from the B to A form. The B form of DNA is the structure stable at high humidity and is the form most prevalent in cells, whereas the A form is stable at lower humidity. The A form is characterized by a tilting of the base pairs relative to the axis of the helix and a change in the position of the sugars. This change in DNA structure appears to be a general consequence of binding zinc fingers and is thought to be involved in the regulation of transcription. The A and B forms of DNA have somewhat different CD spectra, which make it possible to interpret the CD changes associated with zinc finger binding in structural terms.

The interaction of many different proteins and nucleic interactions have been probed with CD (14). These interactions are of great physiological importance, and CD has proved to be a valuable tool for probing the structural changes that occur.

11.9 FLUORESCENCE POLARIZATION

With the introduction of the concept of polarized light, we can introduce an interesting phenomenon observed with fluorescence. Normal fluorescence experiments are not carried out with polarized light so the fluorescence observed also is not polarized. However, if the light used for excitation of fluorescence is polarized, then the fluorescence emitted is also often polarized. The basic idea is not hard to understand: if the electric field wave is oriented in space, it will interact with the electric dipole of the molecule, and this interaction can ultimately lead to the absorption of energy and fluorescence. The polarization of the light emitted will depend on the relative orientation of the electric field and the electric dipole and how much it changes during the fluorescence process. (The polarization of the emitted light is determined by the “transition dipole,” which is generally oriented differently than the electric dipole.) Now let us consider what happens with a population of molecules interacting with polarized light.

The experiment itself is straightforward. A polarizer is put in front of the excitation beam so that molecules are excited with polarized light. A second polarizer is put in front of the detector for the emitted light. Two measurements are made: the

intensity of the emitted light with the axis of the detection polarizer parallel to the polarization axis of the excitation polarizer, $I_{\parallel}$, and the intensity when the two axes are perpendicular, $I_{\perp}$. The polarization, P , is defined as

$$P = \frac{(I_{\parallel} - I_{\perp})}{(I_{\parallel} + I_{\perp})} \quad (11-11)$$

If the emitted light is completely unpolarized, the parallel and perpendicular intensities would be identical and the polarization would be zero. This would happen if the molecules rotate very rapidly during the time that the fluorescence occurs. Since typical fluorescence lifetimes are in the range of 1–100 ns, this would mean that the molecules would have to rotate many times during this relatively short lifetime. This is, in fact, what happens for small molecules where polarization is usually not observed. The opposite extreme is the case when the molecules do not rotate at all during the lifetime of the fluorescence but are randomly oriented in space. This would be the situation, for example, in extremely viscous solvents. The calculation for this situation is somewhat complex, but can be done. For this limit, the polarization is $1/2$. For some macromolecules and ligands bound to macromolecules, the rate of rotation of the fluorescent species is comparable to the fluorescent lifetime. For such cases, the polarization lies between 0 and $1/2$, and measurement of the polarization gives information about the rate of rotation of the fluorophore.

The time dependence of the polarization will provide quantitative information about the rate of rotation of fluorescent molecules. Usually, the data are analyzed in terms of the anisotropy, A , rather than the polarization, although the two are conceptually equivalent. It turns out the anisotropy, defined as

$$A = \frac{(I_{\parallel} - I_{\perp})}{(I_{\parallel} + 2I_{\perp})} \quad (11-12)$$

is more easily related to rotational motion. If the time dependence of the fluorescence is measured for the parallel and perpendicular components, the time dependence of the polarization can be calculated. These are not easy measurements since fluorescence decays in nanoseconds. Theoretical considerations show that for a simple rotation

$$A(t) = A(0) \exp\left(\frac{-t}{\tau_c}\right) \quad (11-13)$$

where t is the time, $A(0)$ is a constant, and τ_c is the rotational correlation time and is related to the rotational diffusion constant, D_{rot} , by $\tau_c = 1/(6D_{\text{rot}})$. For example, the rotational correlation time for a fluorescent derivative of chymotrypsin is 15 ns, which is typical for small proteins. The rotational correlation time becomes longer as the molecular volume of the macromolecule increases, so that it is a direct measure of the size of the macromolecule. The exact relationship is complex, as both molecular weight and shape are important. If more than one mode of rotation is possible, multiple exponential decays may be observed.

Polarization and/or anisotropy provide information about the orientation and rotational freedom of fluorescent molecules. This, in turn, can often provide useful information about biological structures and mechanisms.

11.10 INTEGRATION OF HIV GENOME INTO HOST GENOME

HIV is a retrovirus that is unfortunately well known to everyone. In order for the virus to infect its host, the HIV genome must be integrated into the host genome. This integration is quite complex: it involves processing the HIV DNA, strand transfer, and DNA repair. Integrase is the enzyme responsible for the 3'-processing and strand transfer. These reactions can be monitored *in vitro* with short oligonucleotides that are models for the end of the viral DNA. The catalytic core of the enzyme recognizes a specific sequence of DNA. The self-association of integrase appears to be important for its function. The oligomeric state of integrase bound to viral DNA can be monitored through time-resolved fluorescence anisotropy measurements (16). It is a useful tool for this purpose because very low protein and DNA concentrations can be monitored, and the rotational correlation time is directly related to the size of the macromolecule.

The rotational correlation time for integrase was determined by monitoring the time decay of tryptophan fluorescence. The rotational correlation time decreased from about 90 ns to 20 ns at 25°C as viral-specific DNA, a 21-base sequence, was added. This was attributed to association of integrase with the DNA. The concentration of integrase was only 100 nM in these experiments. Independent studies of integrase established that the rotational correlation time in the absence of DNA was associated with a tetrameric structure of the protein. The actual situation is a bit more complex than this, as some other aggregates are present, but the dominant species is a tetramer, which becomes even more predominant at 37°C. The rotational correlation time of the 21-mer DNA duplex was determined by modifying the DNA with fluorescein. The dominant rotational correlation time is about 2 ns. A much faster rotational correlation time is also observed due to rotation of fluorescein within the DNA. Multiple rotational correlation times are commonly observed so that it is necessary to establish what is being measured on a molecular basis. As the fluorescein-labeled DNA was titrated with integrase, an additional rotational correlation time of about 15 ns was observed. This rotational correlation time is attributed to the complex formed and is consistent with that determined from measurement of the correlation time associated with the tryptophan fluorescence. The rotational correlation time for monomer integrase, molecular weight 32,000, is 16 ns.

These elegant experiments with the integrase–DNA complex, and the many control experiments, demonstrate that the integrase depolymerizes when it binds to DNA. Furthermore, the predominant binding state at 25°C is monomeric, whereas at 37°C a mixture of monomers and dimers is present. This raises the interesting question as to what is the active oligomeric state of integrase. Prior to this study, the prevailing mechanism was thought to involve a multimeric structure of integrase. These results suggest that *in vitro* monomers and dimers may be enzymatically active, although the

in vivo activity, which is considerably more complex, probably involves structures larger than dimers.

11.11 α -KETOGLUTARATE DEHYDROGENASE

The α -ketoglutarate dehydrogenase complex from *E. coli* contains three enzymes that catalyze the overall reaction



The first enzyme decarboxylates the α -ketoglutarate. The intermediate formed with thiamine pyrophosphate transfers the succinyl moiety to lipoamide, and this intermediate is oxidized to form a succinyl-lipoic acid intermediate. The succinyl group is transferred to CoA, and finally the dihydrolipoamide is oxidized. This last reaction involves an enzyme-bound flavin and reduces NAD^+ to NADH. The multienzyme complex has a molecular weight of about 2.5×10^6 and contains 12 copies of the first enzyme, 24 copies of the second, and 12 copies of the third (17). The intermediates in the reaction sequence are bound to lipoamide, that is, lipoic acid covalently attached to a lysine through an amide linkage. The structure of lipoic acid is shown in Figure 11-11. It can exist in an oxidized form with a disulfide at the end of the chain, or in a reduced form with two sulfhydryl groups. The proposed mechanism involves the reduced lipoic acid-bound intermediates moving between the active sites of the three enzymes. Two questions arise: Are the active sites close enough to permit lipoic acid to span the distances between the active sites, and does the lipoic acid rotate between the active sites fast enough for the reaction to proceed efficiently? Answers to both of these questions can be obtained with fluorescence methodology.

Fluorescence resonance energy transfer measurements between the active sites of the three enzymes show that they are about 3 nm apart, approximately the maximum span of a lipoic acid (18). Other experiments showed that succinyl and electron transfer between lipoic acids can occur, so that the potential distance an intermediate can be transferred is longer than a single lipoic acid. Thus, the energy transfer measurements demonstrated that the transfer of intermediates occurs over a relatively long distance and that lipoic acid is a viable intermediate.

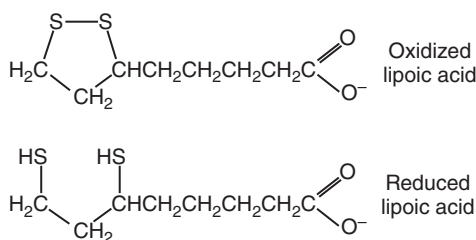


FIGURE 11-11. Structures of oxidized and reduced lipoic acid.

Does the lipoic acid rotate fast enough to serve the postulated transfer function? This is not an easy question to answer because no easy method exists for measuring the rate of rotation. However, fluorescence anisotropy measurements provide some insight into this rate (19). The lipoic acid was labeled with a fluorescent probe on the sulfhydryl that normally carries a reaction intermediate. The fluorescence lifetime of the probe and the dynamic fluorescence anisotropy were then measured. A typical

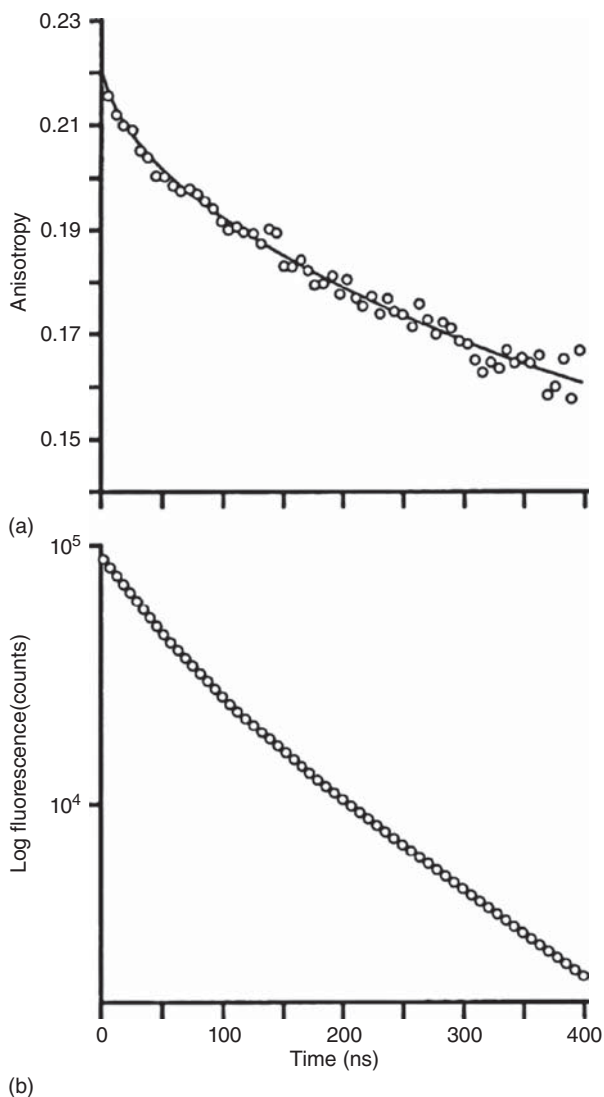


FIGURE 11-12. Time course of the fluorescence (b) and anisotropy (a) decay of pyrene covalently attached to the lipoic acid of *E coli* α -ketoglutarate dehydrogenase. Reprinted in part with permission from D. E. Waskeiwicz and G. G. Hammes, *Biochemistry* **22**, 6489 (1982). © 1982 by American Chemical Society.

time course for the fluorescence and anisotropy is shown in Figure 11-12. As might be expected, the dynamic fluorescence anisotropy is quite complex. Three components are found. One has a very long rotational correlation time and is due to the overall rotation of the very large multienzyme complex. The second component has a rotational correlation time of about 25 ns and is due to local rotation of the fluorescent probe in a hydrophobic environment. Rotational correlation times of this magnitude have been observed often for local motion of ligands bound to macromolecules. The third component has a rotational correlation time of about 350 ns. This can be attributed to rotation of lipoic acid between the catalytic sites. This component of the anisotropy decay is not seen if the enzyme to which the lipoic acid is bound is separated from the other two enzymes. It cannot be ascertained whether the rate-limiting process is the actual rotation or dissociation of the probe from one of the catalytic sites, but in any event these results establish that the rate of rotation is sufficiently fast to support the observed catalytic rate, namely a turnover number of about 130 s^{-1} .

The uses of fluorescence anisotropy to probe the biological mechanisms discussed earlier are intended as illustrative examples of the unique information that can be obtained. They also illustrate the difficulty of interpreting the results that are obtained. As with any physical method, great care must be taken to explore all possible explanations, and careful control experiments are required.

REFERENCES

1. W. B. Gratzer and D. A. Cowburn, *Nature* 222, 426 (1969).
2. N. Sreeramama and R. W. Woody, *J. Mol. Biol.* 242, 497 (1994).
3. N. Berova, K. Nakanishi and R. W. Woody (eds.), *Circular Dichroism: Principles and Applications*, 2nd edition, Wiley, New York, 2000.
4. D. S. Rodgers (ed.) *Circular Dichroism: Theory and Spectroscopy*, Nova Publishers, New York, 2013
5. D. Freifelder, *Physical Biochemistry*, 2nd edition, Freeman, New York, 1982, p. 594.
6. C. A. Sprecher and W. C. Johnson, Jr., *Biopolymers* 16, 2243 (1997).
7. M. Ardhammar, B. Norden, and T. Kurucsev in Ref. 4, pp. 741–768.
8. D. Fornasiero and T. Kurucsev, *J. Phys. Chem.* 85, 613 (1981).
9. J. K. Myers and T. G. Oas, *Annu. Rev. Biochem.* 71, 783 (2002).
10. K. Dill, S. B. Ozken, M. S. Shell, and T. R. Weikl, *Annu. Rev. Biophys.* 37, 289 (2008).
11. S. Altman and L. Kirsebom, *Ribonuclease P in the RNA World* R. F. Gesteland, T. R. Cech, and J. F. Atkins, (eds.), *Cold Spring Harbor Laboratory Press*, Plainview, NY, 1999, pp. 351–380.
12. R. Kole, M. F. Baer, B. C. Stark, and S. Altman, *Cell* 19, 881 (1980).
13. C. H. Henkels, J. C. Kurz, C. A. Fierke, and T. G. Oas, *Biochemistry* 40, 2777 (2001).
14. D. M. Gray in Ref. 3, pp. 769–796.
15. X. Chen, A. Agarwal, and D. P. Giedroc, *Biochemistry* 37, 11152 (1998).
16. E. Duprez, P. Tanc, H. Leh, J.-F. Mouscadet, C. Auclair, M. E. Hawkins, and J.-C. Brochon, *Proc. Natl. Acad. Sci. USA* 98, 10090 (2001).
17. L. J. Reed, *Acc. Chem. Res.* 7, 40 (1974).

18. K. J. Angelides and G. G. Hammes, *Biochemistry* 18, 5531 (1979).
 19. D. E. Waskiewicz and G. G. Hammes, *Biochemistry* 22, 6489 (1982).

PROBLEMS

- 11-1.** The sugar D-mannose can exist as two enantiomers that rotate light in opposite directions. The (+) configuration has a specific rotation of $29.3^\circ \text{ d}^{-1} \text{ g/cc}$ at 589.3 nm, 20°C , whereas the (−) configuration has a specific rotation of $-17.0^\circ \text{ d}^{-1} \text{ g/cc}$ at the same wavelength. When either pure enantiomer is put into water, the optical rotation changes until a specific rotation of $14.2^\circ \text{ d}^{-1} \text{ g/cc}$ is reached. Calculate the ratio of (+) and (−) enantiomers at equilibrium.
- 11-2.** Solutions of RNA and DNA at a concentration of $2.00 \times 10^{-5} \text{ M}$ nucleotides (monomers) have the following differential absorption characteristics in a 1-cm cell:

λ (nm)	$A_L - A_R$, DNA	$A_L - A_R$, RNA
240	-4.40×10^{-4}	0.00
260	0.00	12.0×10^{-4}
280	6.00×10^{-4}	3.20×10^{-4}
300	0.20×10^{-4}	-1.00×10^{-4}

- a. Calculate $\epsilon_L - \epsilon_R$ for the DNA and RNA at each wavelength.
 b. Calculate the molar ellipticity for DNA and RNA at each wavelength.
 c. A mixture of the DNA and RNA has the following differential absorption characteristics.

λ (nm)	$A_L - A_R$
240	-0.53×10^{-4}
260	4.00×10^{-4}
280	1.05×10^{-4}
300	-0.31×10^{-4}

What are the concentrations of DNA and RNA in the mixture?

- d. If the DNA and RNA are hydrolyzed to give the individual nucleotides, will the molar ellipticity increase, decrease, or remain the same?
- 11-3.** A common feature of many DNA binding proteins is the “leucine zipper.” It is two similar sequences of 30–35 amino acids containing multiple leucines. Each sequence forms a right-handed α -helix, and the two helices wrap around

each other to form a left-handed super helix. A number of studies have been carried out of this “coil–coil” structure with model peptides. In one such study, the circular dichroism spectrum at 0°C had a large positive peak at 195 nm and negative minima at 208 and 222 nm. This was interpreted as being the spectrum of a completely helical structure.

- a. At 0°C, the molar ellipticity at 222 nm is $-33,000 \text{ deg cm}^2 \text{ dmol}^{-1}$. At 80°C, the molar ellipticity at 222 nm is essentially zero. Interpret this result.
- b. At 55°C, the molar ellipticity at 222 nm is $-16,000 \text{ deg cm}^2 \text{ dmol}^{-1}$. What percentage of the peptide is α -helical at this temperature?
- c. For some model systems, it is found that the temperature at which the ellipticity approaches zero increases as the concentration of model peptide increases. How would you explain this result?

11-4. Estrogen receptors are ligand-activated transcription factors that mediate the effects of female sex hormone on DNA transcription. The interaction of estrogen receptors with a specific DNA fragment that binds the receptor and estradiol has been studied with circular dichroism [N. Greenfield, V. Vijayanathan, T. J. Thomas, M. A. Gallo, and T. Thomas, *Biochemistry* **40**, 6646 (2001)]. The circular dichroism spectrum was analyzed to show that the receptor is approximately 75% α -helical, 3% β -sheet, 10% turns, and 12% random coil. The following data were obtained for the ellipticity at 222 nm in the presence and absence of the DNA and estradiol.

Ligand	$[\theta]_{222}, \text{deg cm}^2 \text{ dmol}^{-1}$
None	$-25,000$
5 μM DNA	$-32,000$
5 μM Estradiol	$-22,000$

- a. Assume that the specific rotation at 222 nm is due only to the α -helix and that the receptor is saturated with the ligand. Indicate what is happening to the structure of the receptor and calculate the percentage of helix when the ligand is bound.
- b. When the temperature is raised, the ellipticity at 222 nm approaches zero. Explain what is happening to the structure of the receptor.
- c. The temperature at which the difference between the initial and final ellipticities has reached half of its value, T_m , is 38.0°C in the absence of ligands. It is 43.6°C and 46.1°C in the presence of 5 μM estradiol and 5 μM of the specific binding DNA, respectively. Explain these results.
- d. In the presence of 750 nM estradiol, T_m is 40.8°C, Explain this result and estimate the dissociation constant for the binding of estradiol to its receptor protein with the assumption that the concentration of estradiol is much greater than the receptor concentration. At 40.8°C, the ratio of denatured

protein to native protein was determined to be 1.20 in the absence of ligands. (This is a hypothetical result as this result was not reported in the publication.)

- 11-5.** Adenosine monophosphate has an extinction coefficient of about $15,000 \text{ cm}^{-1} \text{ M}^{-1}$ at 260 nm, 0°C . The value of $\epsilon_L - \epsilon_R$ at 260 nm is about $2.00 \text{ cm}^{-1} \text{ M}^{-1}$. The corresponding value of $\epsilon_L - \epsilon_R$ for a polymer of polyrA is $-17.0 \text{ cm}^{-1} \text{ M}^{-1}$ at 0°C , and its magnitude decreases greatly as the temperature is raised. (The monomer concentration was used in calculating this number.)
- Explain the reason for the large difference in the differential extinction coefficient between AMP and polyrA.
 - Why does the magnitude of the differential extinction coefficient decrease as the temperature is increased?
 - Calculate the observed ellipticity in degrees for a 10^{-4} M solution of AMP and for a 10^{-4} M solution of polyrAMP (monomer concentration) at 260 nm, 0°C .
 - Calculate the absorbance of a 10^{-4} M solution of AMP. Would the absorbance of the polyrA at the same concentration (0°C) be greater than, less than, or the same as the ATP solution?
- 11-6.** The system for transport of mannitol across the *Escherichia coli* membrane involves a membrane-bound enzyme that couples phosphorylation of the sugar to its translocation through the membrane. Fluorescence lifetime and anisotropy measurements have been carried out of the tryptophans in this enzyme dissolved in detergent micelles. [D. Dijkstra, J. Broos, A. J. W. G. Visser, A. van Hoek, and G. T. Robillard, *Biochemistry* **36**, 4860 (1997)]. The fluorescence anisotropy decay of one of these tryptophans has two components. The faster component, with a rotational correlation time of about 1 ns, is due to local motion of the tryptophan. The decay of anisotropy for the slower component can be approximated by the following data.

Anisotropy	Time (ns)
0.290	0.00
0.245	5.00
0.208	10.00
0.149	20.00
0.126	25.00
0.106	30.00
0.090	35.00

- Calculate the rotational correlation time.

- b.** The expected correlation time for the known molecular weight of the macromolecule is 120–140 ns. How do you explain the difference between the calculated and observed correlation times?
- c.** When the protein is phosphorylated, the rotational correlation time increases to 51 ns. What might be the cause of this change?

Vibrations in Macromolecules

12.1 INTRODUCTION

Thus far, we have considered only spectral phenomena in the visible and ultraviolet regions of the spectrum. In this spectral region, the interaction between light and molecules causes transitions between energy levels of the electrons and alters the populations of electronic energy levels. Molecules also have characteristic vibrational motions that are influenced by much longer wavelengths, typically in the far red or infrared. Since the energy associated with a given wavelength, λ , is hc/λ , this means that the difference in energy between levels is much smaller than that for the electronic energy scaffold. (Recall that h is Planck's constant and c is the speed of light.)

In principle, some type of coupling might be expected when transitions occur between electronic and vibrational energy levels. However, this is not the case because electronic transitions occur much more rapidly than the time scale of nuclear motions—nuclei are quite sluggish compared to electrons (10^{-16} s vs. 10^{-13} s). Thus, the nuclei can be assumed to be stationary during an electronic transition. This is called the Franck–Condon principle. The Franck–Condon principle is why electronic transitions between vibrational energy levels of different electronic energy states can be drawn as straight lines (Fig. 10-11).

An analysis of molecular vibrations utilizing quantum mechanics has been presented in Chapter 7. In this treatment, vibrational motion has been approximated by a harmonic oscillator, that is, essentially atoms connected by springs. Before considering the experimental aspects of vibrational spectroscopy, the following salient points should be recalled.

A detailed analysis of all of the $3N$ normal mode vibrations of a macromolecule is not possible. Nevertheless, in some limiting cases, these normal modes are dominated by the movement of a single or a restricted number of chemical bonds, for example, a double bond between carbon and oxygen. In these cases, the normal mode has the same signature frequency in different molecules.

The energy levels for the harmonic oscillator are given by the relationship

$$E = \left(n + \frac{1}{2}\right) h\nu \quad (12-1)$$

where E is the energy, ν is the frequency, and n is the vibrational quantum number, which is zero or a positive integer.

The characteristic frequency of the harmonic oscillator bond motion is

$$\nu = \left(\frac{1}{2\pi} \right) \left(\frac{k}{\mu} \right)^{1/2} \quad (12-2)$$

where k is the spring constant, essentially a measure of the “stiffness” of the spring, and μ is the reduced mass of the system, which is a weighted average of the masses of the atoms involved in the normal mode. As shown by Eq. (12-1), the energy levels are equally spaced with splittings of $h\nu$. The larger the force constant k , the stiffer the spring and the greater the splittings between the vibrational energy levels. For the harmonic oscillator, quantum mechanical considerations only permit changes in n of ± 1 .

As discussed in Chapter 7, for real molecules the vibrational motion deviates from a simple harmonic model at high energies, and the energy levels become more closely spaced at the top of the potential well. However, for most molecules at room temperature, the harmonic oscillator is a good approximation for vibrational motions.

12.2 INFRARED SPECTROSCOPY

The difference in energy between different vibrational energy levels is about 100 J/mol. This corresponds to light in the infrared region, $\lambda \sim 1$ mm. Wave numbers, $\bar{\nu}$, are usually used when discussing vibrational spectra, rather than wavelength. The wave number is simply the reciprocal of the wavelength. Quantum mechanical calculations indicate that there are selection rules that govern the transitions between vibrational energy levels. If absorption of light is to occur, the vibration must cause a change in the electric dipole moment. Recall that a dipole moment is simply a measure of the balance of charges within a molecule. For example, HCl has a permanent dipole moment because H has a partial positive charge and Cl has a partial negative charge. On the other hand, H_2 does not have a permanent dipole moment because no net imbalance in charge is present. From a practical standpoint, it is important to note that water has a dipole moment and absorbs light in the infrared very strongly. Consequently, it is very difficult to measure the infrared absorption of macromolecules in water as the absorption–emission of the molecule is significantly obscured by that of water. Sometimes, a combination of H_2O and D_2O can be used to partially circumvent this problem, but it severely restricts the use of infrared spectroscopy for biological systems. In some cases, dried and/or hydrated films are used, although this is far from the conditions in the milieu of biology.

Two experimental methods are available for measuring infrared spectra. One is essentially the same as the visible ultraviolet spectrometer previously discussed. Infrared light is passed through the sample, and the absorption is measured. The only difference is the light source, typically a glowing wire, and the detector, typically a thermocouple. This method has been largely supplanted by Fourier transform

methods. These methods were briefly described in Chapter 7. Basically, with Fourier transform infrared measurements, a beam of light is split into two, with only half of the light going through the sample. The difference in phase of the two waves creates constructive and/or destructive interference and is a measure of the sample absorption. The waves are rapidly scanned over a specific wavelength region of the spectra, and multiple scans are averaged to create the final spectrum. This method is more sensitive than the conventional dispersion spectrometer.

12.3 RAMAN SPECTROSCOPY

Another method exists for studying transitions between vibrational energy levels that uses a concept not yet discussed. In addition to absorbing light, samples also scatter light. The amount of scattered light is a maximum 90° to the direction of the incident beam. Most of the scattered light is at the same frequency as that of the incident light. This is called Rayleigh scattering. At the molecular level, the electric field of the light perturbs the electron distribution, but no transitions between energy levels occur so that the molecule immediately returns to its unperturbed state. This scattering is inversely proportional to the fourth power of the wavelength so the scattering is much greater at shorter wavelengths. This is essentially why the sky is blue: the shorter wavelengths of the visible spectrum (blue) are scattered more than the longer wavelengths (red). Rayleigh scattering is observed at all wavelengths. The intensity of the scattered light is related to the *polarizability* of the molecule, that is, to how easily electronic charge can be shifted within the molecule to make it more polar. A small number of the molecules return to a different vibrational energy level after scattering. The vibrational energy level can be either higher or lower than the initial state. As a result of this change in energy level, some of the scattered light will be at a slightly lower or higher frequency than the incident light. This is called *Raman* scattering, after the Indian scientist who discovered the phenomenon.

A typical setup for measuring Raman scattering is shown in Figure 12-1. A very intense light source is needed to observe Raman scattering because only a very small amount of the scattered light displays a change in frequency. The advent of lasers has permitted this to be done routinely with visible light and in the ultraviolet with rather expensive lasers. Prior to lasers, very large arc lamps were utilized. In addition, very high concentrations are required. However, Raman scattering does not require the light to be at a wavelength comparable to the energy of vibrational transitions because it is a *scattering* phenomenon, as contrasted to absorption spectroscopy. The Rayleigh line is very intense, but much less intense scattering can be detected at a lower frequency than the incident light. This is because the incident light was used to promote the molecule to a higher vibrational energy level. These are called Stokes lines. An even smaller fraction of the scattered light occurs at a higher frequency than the incident light because energy is added to the incident light by the movement of the molecule to a lower vibrational energy level. These are called anti-Stokes lines. Most molecules are found in their ground vibrational energy level at room temperature so that the observation of anti-Stokes lines is rare.

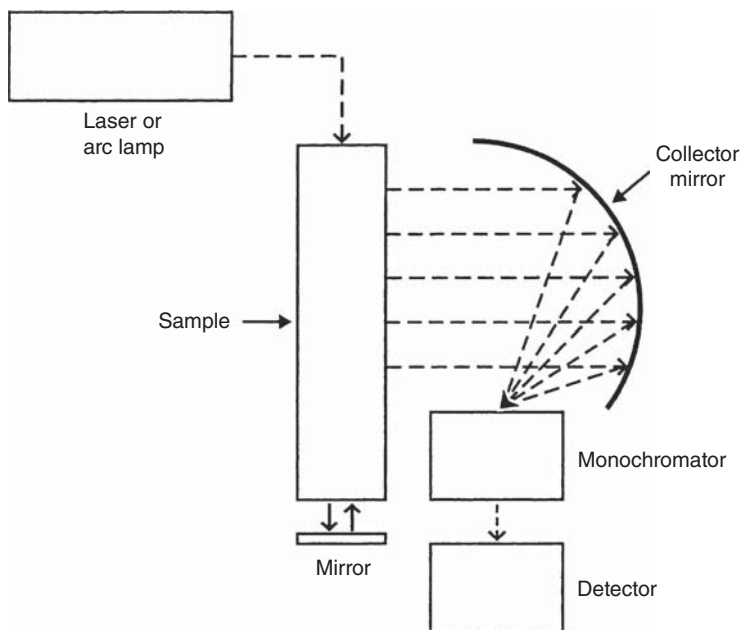


FIGURE 12-1. Schematic diagram of an apparatus used for Raman spectroscopy. The mirror at the end of the sample cell is to put the light through the sample a second time, effectively increasing the path length of the cell, and the collector mirror is designed to collect as much of the scattered light 90° to the source light beam as possible.

Raman spectroscopy has two major advantages over infrared spectroscopy for studying transitions between vibrational energy levels. First, a permanent dipole moment is not required. It is only necessary for the polarizability of the molecule to change between different vibrational energy levels. Second, visible light can be used rather than infrared light so that Raman spectra can be readily obtained in water. They can also be obtained in crystals and films. As previously indicated, the primary disadvantages are that because the intensity of the Raman lines is very weak, intense light sources and high concentrations of the molecule of interest are needed. Raman and infrared spectroscopy should be regarded as complementary. Since infrared spectroscopy is dependent on the permanent dipole moment and Raman spectroscopy on the polarizability, usually (but not always) a vibrational transition is observed either in the infrared or in Raman scattering, but not in both.

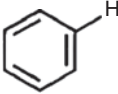
Thus far in our discussion, the implicit assumption is that the wavelength of the exciting light for Raman spectroscopy is not near an absorption band. If the wavelength coincides with the absorption band for an electronic transition, a large increase occurs in the intensity of the Raman spectrum. Basically, this is related to the fact that an electronic absorption band is the superposition of many different vibrational modes. This can sometimes be seen directly in the electronic absorption spectrum by fine structure in the peaks. Essentially, the alteration of the population of vibrational

energy levels within the electronic absorption band is responsible for the enhanced intensity of the Raman spectrum. Determining the spectrum within the electronic absorption band is called *resonance Raman spectroscopy*. The twofold advantage of resonance Raman spectroscopy is that lower concentrations can be used and the spectrum is simplified because only the intensified lines are observed. At the present time, resonance Raman measurements are made primarily in the visible region of the spectrum. In biological systems, this means chromophores such as hemes and retinal must be present.

12.4 STRUCTURE DETERMINATION WITH VIBRATIONAL SPECTROSCOPY

Infrared spectroscopy has been a useful tool for the determination of the structure of organic molecules for many years. This is because specific types of bonds and/or chemical groups have characteristic vibrational frequencies. Some of these group frequencies are given in Table 12-1. Infrared spectra are often referred to as

TABLE 12-1. Group Frequencies in the Infrared Region^a

Chemical group frequency (cm ⁻¹)	
-CH ₃	1460
-CH ₂ -	2930
	2860
	1470
C-H	3300
-C-C-	1165
-C=O	1730
-C-H (in CH ₃)	2960
	2870
-C-H (in CHO)	2870
	2720
	3060
	
-CN	2250
-O-O-	1200-1100
-OH	3600
-NH ₂	3400
=CH ₂	3030
-SH	2580
-C=N-	1600
C-Cl	725
C=S	1100

^aReproduced with permission from D. Sheehan, *Physical Biochemistry*, John Wiley and Sons, 2000, p. 98.

TABLE 12-2. Protein Secondary Structure Determined by Infrared (IR) and Circular Dichroism (CD) Spectra and X-Ray Crystallography

Protein	Secondary structure (%)				Method
	α -Helix	β -Sheet	Turn	Random	
Hemoglobin	78	12	10	^a	IR
	87	0	7	6	X-ray
	68–75	1–4	15–20	9–16	CD
Myoglobin	85	7	8	^a	IR
	85	0	8	7	X-ray
	67–86	0–13	0–6	11–30	CD
Lysozyme	40	19	27	14	IR
	45	19	23	13	X-ray
	29–45	11–39	8–26	8–60	CD
Cytochrome <i>c</i> (oxidized)	42	21	25	12	IR
	48	10	17	25	X-ray
	27–46	0–9	15–28	28–41	CD
α -Chymotrypsin	9	47	30	14	IR
	8	50	27	15	X-ray
	8–15	10–53	2–22	38–70	CD
Trypsin	9	44	38	9	IR
	9	56	24	11	X-ray
Ribonuclease A	15	40	36	9	IR
	23	46	21	10	X-ray
	12–30	21–44	11–22	19–50	CD
Alcohol dehydrogenase	18	45	23	14	IR
	29	40	19	12	X-ray
Concanavalin A	8	58	26	8	IR
	3	60	22	15	X-ray
	3–25	41–49	15–27	9–36	CD
Immunoglobulin G	3	64	28	5	IR
	3	67	18	12	X-ray
Major histocompatibility complex antigen A2	17	41	28	14	IR
	20	42			X-ray
	8–13	74–77			CD
β_2 -Macroglobulin	6	52	33	9	IR
	0	48			X-ray
	0	59			CD

Reproduced with permission from A. Dong, P. Huang, and W. S. Caughey, *Biochemistry* **29**, 3303 (1990). © 1990 American Chemical Society.

^aThe band due to random structure appears as a shoulder on the α -helix band and is too small to be separated from α -helix structure. The random structure is estimated at <5% and is included in the α -helix value.

“fingerprints” for the molecule, and large compilations of data are available. In the case of biological molecules, changes in the group frequencies can be used to derive information about the secondary structure of the molecules (1).

The carbonyl of the amide bond in proteins is particularly useful for the determination of secondary structure (2,3). The stretching normal mode, amide I mode, of the carbonyl has been shown to have a specific frequency associated with α -helices, β -sheets, and other characteristic structures. (Strictly speaking, this normal mode is not just the stretching of the carbonyl. It also involves some bending of the C–N–H angle.) This was ascertained by the study of model peptides for which precise measurements and theoretical calculations of normal modes could be carried out. The approximate wave numbers corresponding to the three common structures found in proteins are as follows: α -helix, 1650 cm^{-1} ; β -sheet, 1632 and 1685 cm^{-1} ; and random coil, 1658 cm^{-1} .

For proteins, a more empirical approach has been adopted. The most successful approach uses known structures to calibrate the vibrational frequency measurements. The vibrational spectrum of the amide bond for a protein is complex because of the many amide bonds present in multiple environments. Nevertheless, the spectrum can be deconvoluted to provide information about the amount and types of secondary structures present. The assumption usually made is that the observed vibrational frequency is a linear combination of the frequencies associated with the various secondary structures that are present, with each specific frequency weighted by the percent of a given structure. Some typical results are shown in Table 12-2 and compared with the known structure of the protein and CD estimates of secondary structure (4). Raman spectroscopy has proven particularly useful for studying secondary structure of proteins because water solutions can be used, and the amide I mode has a spectral band that is well isolated from other protein bands ($1630\text{--}1700\text{ cm}^{-1}$). The amide I band is the most common vibrational frequency used as an indicator of secondary structure (cf. Ref. 5), but other bands and their relation to structure have been identified (3).

The structure of nucleic acids can also be investigated with vibrational spectroscopy. The base vibrations and the groups involved in hydrogen bonding are particularly sensitive to the secondary structure of nucleic acids. Most of the work with nucleic acids has involved Raman spectroscopy (3). For example, the melting of DNA and RNA structures can be readily followed, and the specific vibrational frequencies provide molecular details about the melting process. The different types of helices formed by DNA can be distinguished. Standards are established with known structures and then used to determine the structures of unknown samples, that is, the amount of various types of structures present.

12.5 RESONANCE RAMAN SPECTROSCOPY

Heme proteins have been extensively studied with resonance Raman spectroscopy (6). They are very prevalent in nature and have a very intense absorbance in the visible region of the spectrum due to the porphyrin ring structure. Excellent Raman

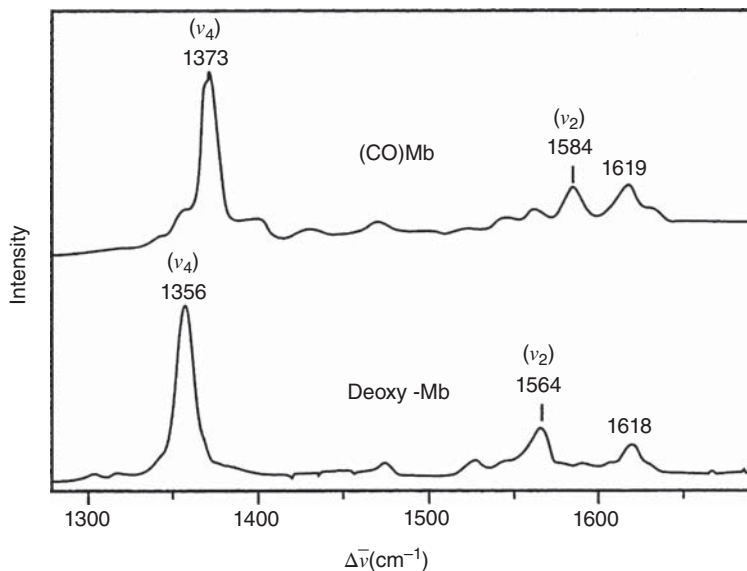


FIGURE 12-2. Comparison between resonance Raman spectra for carbonmonoxymyoglobin [(CO)Mb] and deoxymyoglobin at room temperature. The excitation was in the wavelength region corresponding to the heme absorption. Note that the two species can be readily distinguished by their Raman spectra. Reprinted from T. Spiro and R. S. Czernuszewicz, *Resonance Raman spectroscopy of metalloproteins, Methods Enzymol.* **246**, 416 (1995). © 1995, with permission from Elsevier.

spectra can be obtained at very low concentrations, often in the micromolar range. The vibrational spectra obtained are characteristic of the porphyrin ring structure. However, the highest frequency normal modes, $1350\text{--}1650\text{ cm}^{-1}$, are very sensitive to the state of the Fe atom that is bound to the porphyrin. Thus, these frequencies can establish the spin state and coordination states of the Fe in both of its oxidation states. The binding of ligands to the Fe and the distortion of the porphyrin skeleton can also be detected. As an example, the high-frequency region of the spectrum for deoxy- and CO-myoglobin is shown in Figure 12-2 (6). Myoglobin is a protein that is used for oxygen transport in some organisms, a function carried out by hemoglobin in humans. The frequency shifts in the vibrational spectra between these two forms of myoglobin can be readily discerned.

The details of ligand binding in hemoglobin have been frequent targets of Raman spectroscopy (cf. Ref. 7). As an example, consider the study of a naturally occurring mutant labeled hemoglobin M Iwate (8). This mutant has a tyrosine substituted for a histidine in the α chain. The resonance Raman spectrum of the tyrosine was determined by the scattering of 244 nm radiation, and the resonance Raman of the heme was determined using scattering at 406 and 441 nm. The ultraviolet Raman spectra for normal hemoglobin and hemoglobin M Iwate are shown in Figure 12-3, along with the difference spectrum. The vibrations seen in the difference spectrum are attributed

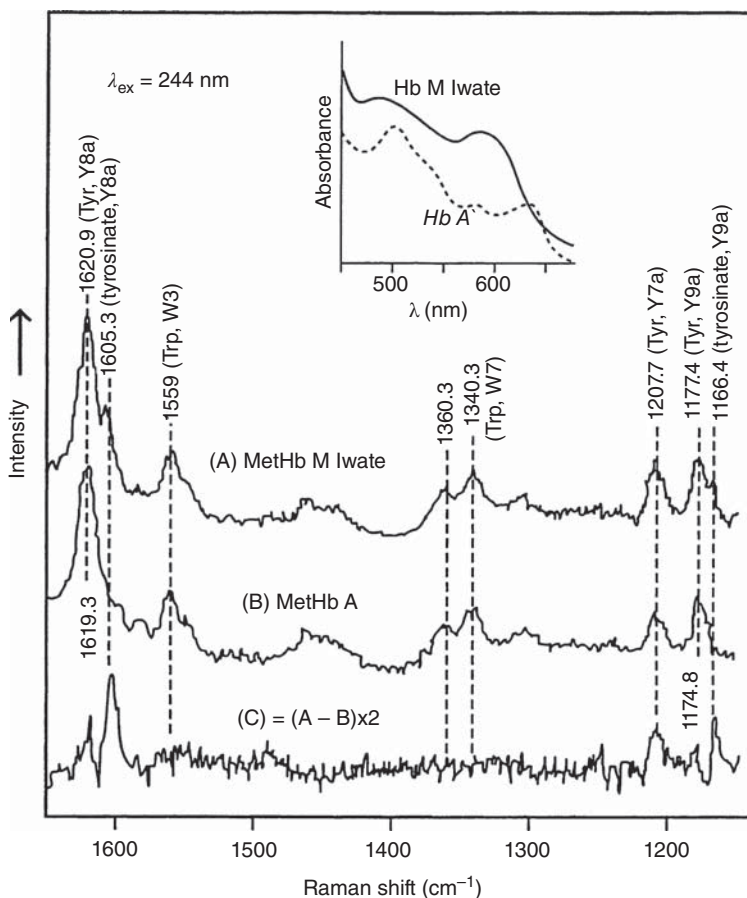


FIGURE 12-3. Resonance Raman spectra of two different hemoglobin (Hb) variants with excitation at 244 nm. The spectra of metHb M Iwate (A), metHb A (B), and their difference, expanded twofold (C), are shown. The inset shows the absorption spectra of the two Hbs in the visible region. Reprinted in part with permission from M. Nagai, M. Aki, R. Li., Y. Jin, H. Sakai, S. Nagatomo, and T. Kitagawa, *Biochemistry* **39**, 13093 (2000). © 2000 by American Chemical Society.

to the extra tyrosine in the mutant. By comparison with known compounds, it was shown that the tyrosine is deprotonated in the mutant and is bound to the Fe(III) heme in the α -subunit.

The effect of this binding on the overall heme structure was determined from the resonance Raman spectra in the visible range. Furthermore, reduction of the Fe(III) to Fe(II) resulted in the elimination of the tyrosinate vibrations, and it could be shown that a histidine residue was coordinated to the Fe(II) heme. This study is a good example of how vibrational spectroscopy can be used to elucidate the detailed structure of ligand–protein interactions.

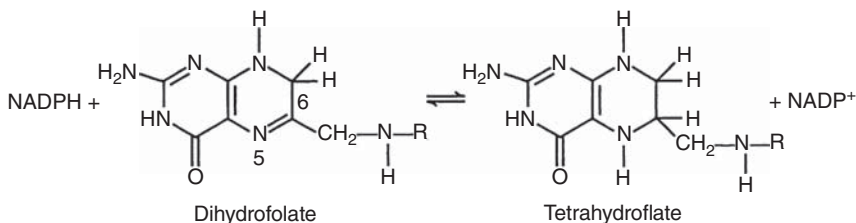


FIGURE 12-4. The reaction catalyzed by the enzyme dihydrofolate reductase. The NADPH reduces dihydrofolate to tetrahydrofolate with the production of NADP^+ . Note that a proton, obtained from the solvent, is needed on the left-hand side of the equation to give a balanced chemical reaction.

12.6 STRUCTURE OF ENZYME-SUBSTRATE COMPLEXES

One of the advantages of vibrational spectroscopy is that subtle changes in the electronic environment can be related to structural changes that may be difficult to observe by more direct methods, such as X-ray crystallography. This feature has been used to elucidate the catalytic process for enzymatic reactions (cf. Ref. 9). As an example, consider the enzyme dihydrofolate reductase (DHFR), which has been discussed earlier (Chapter 10). The reaction catalyzed converts 7,8-dihydrofolate (H_2 folate) to 5,6,7,8-tetrahydrofolate by catalyzing the transfer of a hydride ion from NADPH to C6 and addition of a solvent proton to N5 (Fig. 12-4).

An important mechanistic question centers on the protonation state of N5 of dihydrofolate. Is the nitrogen protonated prior to the hydride transfer or is the proton added after the hydride transfer occurs? Raman spectroscopy showed that an N5–C6 stretching vibration changed upon protonation of N5 from 1650 cm^{-1} to 1675 cm^{-1} (10). This was established by the use of model compounds, isotope effects on the vibrational spectrum, and theoretical calculations. On the basis of these extensive studies of model systems, N5 was found to be unprotonated in the DHFR– H_2 folate complex at neutral pH, but it was protonated in the DHFR– H_2 folate– NADP^+ complex. A pH titration indicated that the $\text{p}K_a$ of N5 is 6.5 in this complex, as compared to 2.6 when the substrate is not bound to the enzyme. In the actual reaction, NADPH is present rather than NADP^+ so that the implicit assumption is that the complex studied is a good model for the catalytic reaction. Thus, N5 can be protonated much more readily in the enzyme– NADP^+ complex.

From the point of view of understanding enzyme catalysis, this result means that the population of the N5 protonated substrate is four orders of magnitude larger in the environment of the enzyme. The positive charge on the N5 would also make C6 more positive and would presumably make the hydride transfer reaction much faster than with unprotonated N5. Much of the catalytic effect of the enzyme, therefore, apparently is due to providing an environment that stabilizes the protonated substrate prior to the hydride transfer. A structural explanation for this stabilization is not obvious. The N5 is in a hydrophobic region of the protein, and no negative charges are

conveniently close that might stabilize the protonated species. However, conformational changes of the protein have been established as part of the overall mechanism of action of the enzyme, and these changes may be the source of the stabilization. In any event, Raman spectroscopy has provided unique insight into the catalytic process.

12.7 CONCLUSION

Vibrational spectroscopy can also be utilized to monitor complex biological processes. For example, a recent study monitored oxygen evolution during photosynthesis (11). The oxygen is generated by the oxidation of water at a Mn_4CaO_5 cluster site. The oxygen evolution was initiated with a flash lamp, and the time courses of the absorption changes at 1400 and 2500 cm^{-1} were followed as a function of time with Fourier transform infrared spectroscopy. These spectral bands are associated with the reactions and/or interaction changes of carboxylate groups and the changes in polarizability of strong hydrogen bonds, respectively. A large number of reaction intermediates are observed following the flash. The time course of changes in the infrared spectra provided experimental evidence that significant proton rearrangements occur, most likely reflecting the release of a proton from the catalytic site, and that a transient state forms before the oxidation of the Mn_4CaO_5 cluster that leads to oxygen production. Other proton movements were also detected. These results were used to develop a detailed reaction mechanism for the oxygen production. Although vibrational spectroscopy has not been used as extensively in biological systems as ultraviolet–visible absorption and fluorescence spectroscopy, it can sometimes provide unique and important information.

REFERENCES

1. K. Sauer, (ed.), *Methods Enzymol.* 246 (1995).
2. R. Siebert, *Methods Enzymol.* 246, 501 (1995).
3. W. L. Peticolas, *Methods Enzymol.* 246, 389 (1995).
4. A. Dong, P. Huang, and W. S. Caughey, *Biochemistry* 29, 3303 (1990).
5. W. K. Surewicz, H. H. Mantsch, and D. Chapman, *Biochemistry* 32, 389 (1993).
6. T. Spiro and R. S. Czernuszewicz, *Methods Enzymol.* 246, 416 (1995).
7. D. L. Rousseau and J. M. Friedman, *Biological Application of Raman Spectroscopy*, Vol. 3 (T. G. Spiro, ed.), Wiley, New York, 1988 p. 133.
8. M. Nagai, M. Aki, R. Li, Y. Jin, H. Sakai, S. Nagatomo, and T. Kitagawa, *Biochemistry* 39, 13093 (2000).
9. H. Deng and R. Callender, *Infrared and Raman Spectroscopy of Biological Materials* (H.-U. Gremlich and B. Yan, eds.), Marcel Dekker, Inc., New York, 2001, pp. 477–514.
10. Y.-Q. Chen, J. Kraut, R. E. Blakley, and R. Callender, *Biochemistry* 33, 7021 (1994).
11. T. Noguchi, H. S. Suzuki, M. Tsuno, M. Sugiura, and C. Kato, *Biochemistry* 51, 3205 (2012).

PROBLEMS

- 12-1.** The first four vibrational energy levels ($n=0, 1, 2, 3$) for NaI are: 283.67, 848.80, 1411.0, and 1970×10^{-23} J/molecule.
- Prove that these energy levels are not consistent with the harmonic oscillator molecule.
 - What is the wavelength of light that would be emitted when the molecule drops from the $n=1$ to $n=0$ level? From the $n=3$ to $n=2$ level?
 - For an anharmonic oscillator, transitions between energy levels can occur other than changes in n of ± 1 . What wavelength is the emitted light when the molecule moves from $n=2$ to $n=0$?
- 12-2.** From infrared studies of model proteins, the wave numbers for the amide II band (due to N—H deformation in the peptide bond) are $1540\text{--}1550\text{ cm}^{-1}$ for the α -helix, $1520\text{--}1525\text{ cm}^{-1}$ for β -sheet, and $<1520\text{ cm}^{-1}$ for “random coils.”
- Calculate the zero-point energy for these three vibration energy levels.
 - Explain why the three fundamental wave numbers differ in terms of the protein structure.
 - For polyglutamic acid, the amide II band is at a wave number of about 1545 cm^{-1} at low pH (pH 4). As the pH increases ($>\text{pH } 9$), the wave number decreases to below 1520 cm^{-1} . Explain this result.

Principles of Nuclear Magnetic Resonance and Electron Spin Resonance

13.1 INTRODUCTION

In this chapter, we consider the interaction of molecules with radiation when molecules are placed into a strong magnetic field. The fundamental properties of atoms that are important for this discussion are the nuclear spin, for nuclear magnetic resonance (NMR), and the electron spin, for electron spin resonance (ESR) or, equivalently, electron paramagnetic resonance (EPR). Strictly speaking, the concept of spin can be rigorously defined only by the use of quantum mechanics. However, we will use a semiclassical approach in which spin in the nucleus or in an electron can be represented as a charge moving in a circular path. This movement creates a magnetic dipole that can be thought of as a bar magnet. In the absence of a magnetic field, the magnetic dipole is oriented randomly, and only one energy level is associated with the electron or nucleus. But in the presence of a magnetic field, the magnetic dipoles (or bar magnets) tend to be oriented either in the direction of the field or opposed to it, thus creating multiple energy states. Application of quantum mechanics to this situation indicates that the orientation of the magnetic dipole and the energy states are quantized with characteristic quantum numbers.

For electrons, the spin quantum number, S , is $1/2$, and the two spin states are $+1/2$ and $-1/2$, represented as the familiar arrows pointed up or down. Neutrons and protons also have a spin quantum number of $1/2$ so that the nucleus of an atom has a characteristic spin quantum number, I . Simple rules exist for determining the nuclear spin quantum number. Nuclei with an even mass number and even charge number have no nuclear spin ($I = 0$). Nuclei with an odd mass number have a half integral spin ($I = 1/2, 3/2, 5/2$, etc.). Finally, nuclei with an even mass number and odd charge number have integral spin ($I = 1, 2$, etc.).

The number of energy levels associated with spin is determined by the spin quantum number. For an electron, the number of energy levels is $2S + 1$, or 2. In the absence of a magnetic field, no distinction can be made between these two quantum

states: They have the same energy, that is, they are degenerate. In the presence of a magnetic field, however, the alignment of the magnetic moments with and against the magnetic field creates two distinct energy levels. The hydrogen nucleus also has a spin quantum number, I , of $1/2$, with two possible orientations of its magnetic dipole (or magnetic moment) in a magnetic field and two energy levels ($2I + 1$). On the other hand, ^{23}Na has $I = 3/2$ and four energy levels in the presence of a magnetic field. In general, the quantum states in the presence of a magnetic field are characterized by quantum numbers ranging from I to $-I$ in integral steps. Thus, for sodium, these quantum numbers are $3/2$, $1/2$, $-1/2$, and $-3/2$. Although the average orientation of the nuclear spin state of a proton is either aligned with the magnetic field or against it, the conservation of angular momentum requires the magnetic moment associated with the orientation to rotate about the direction of the field. This is analogous to a top spinning on its axis and rotating around a vertical line due to gravity.

A quantum mechanical treatment of nuclear spins in a magnetic field of a strength H provides an explicit equation for the energy levels, E :

$$E = -g_N\beta_N H m_I = -\left(\frac{h}{2\pi}\right)\gamma H m_I \quad (13-1)$$

In this equation, β_N is the nuclear magneton and is a universal constant calculated from the properties of nuclei: $\beta_N = 5.051 \times 10^{-27}$ Joules/Tesla. The nuclear g factor, g_N , is a constant, but it is different for each atom, and m_I is the spin quantum number characterizing the orientation of the magnetic moment in the magnetic field (I to $-I$ in integral steps). This equation also defines the gyromagnetic ratio, γ , which is a frequently used constant. Values of the nuclear spin quantum number, I , and γ for some nuclei of biological interest are presented in Table 13-1.

The dependence of the energy on magnetic field for atoms with a total nuclear spin quantum number of $1/2$ is shown in Figure 13-1. The difference in energy between the quantized levels increases as the magnetic field increases. Also shown is the precession of the magnetic moment about the field direction for the two possible orientations.

The energy difference between the energy levels can be calculated, again with the assistance of quantum mechanics, by the requirement that the change in quantum

TABLE 13-1. Magnetic Properties of Selected Nuclei

Isotope	Spin	$10^7\gamma(\text{T}^{-1} \text{ S}^{-1})$	Natural abundance
^1H	$1/2$	26.75	99.98
$^2\text{H}(\text{D})$	1	4.11	0.0156
^{13}C	$1/2$	6.73	1.108
^{14}N	1	1.93	99.63
^{15}N	$1/2$	-2.75	0.37
^{19}F	$1/2$	25.18	100.0
^{31}P	$1/2$	10.84	100.0
^{17}O	$5/2$	-3.63	0.037

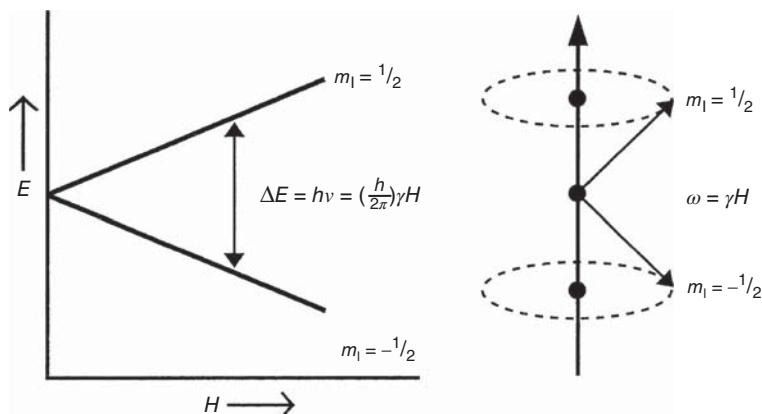


FIGURE 13-1. On the left, a schematic plot of the energy versus the magnetic field is shown for a nuclear or electron spin with a quantum number of $1/2$. The frequency of the radiation emitted/absorbed for a transition between the two energy states created by the magnetic field is dependent on the strength of the magnetic field (Eq. (13-2)). On the right, the magnetic dipole associated with the nuclear spin is shown precessing around the direction of the magnetic field at an angular frequency ω . The two orientations, up and down, correspond to the two energy levels in the diagram on the left.

number for allowed transitions between energy levels is 1. In this case,

$$\Delta E = g_N \beta_N H = \left(\frac{h}{2\pi} \right) \gamma H \quad (13-2)$$

With this equation we can calculate the frequency, ν , associated with transitions between energy levels since $\Delta E = h\nu$. The magnetic field of modern instruments varies from about 7 to 19 Tesla. For a magnetic field of 11.75 Tesla, the frequency associated with transitions between the energy levels of protons is 500 MHz. This is in the radio frequency range. The actual energy difference between levels is quite small, only 3.4×10^{-25} J/proton. Because the energy difference is small, the actual population difference between energy levels is also very small at room temperature: the ratio of populations in the two states is $\exp[-\Delta E/(k_B T)] = 0.99993$ for an 11.75 Tesla magnetic field at 37°C . For this same magnetic field, the resonant frequency for ^{13}C is 130 MHz.

The quantum mechanical calculation of the energy levels for an unpaired electron in a magnetic field is quite similar to that for nuclear spin, except that the electron always has spin $1/2$. The energy difference between levels is

$$\Delta E = g_S \beta_S H \quad (13-3)$$

where $g_S = 2.0023$ and $\beta_S = 9.274 \times 10^{-24}$ Joules/Tesla for electrons. Because the Bohr magneton is much larger for an electron than a proton, the energy levels are further apart, and the resonance energy frequency is much larger. (β_S and β_N differ by the ratio of the mass of the proton to the mass of the electron, 1836.) For a 1

Tesla field, the resonant frequency is $28,000 \text{ MHz} = 28 \text{ GHz}$. This frequency is in the microwave region, and quite different experimental techniques are required for ESR and NMR.

Finally, we return to the precession of the magnetic moment about the direction of the magnetic field, as depicted in Figure 13-1. The angular precession frequency, that is, how fast the magnetic moment is rotating about the vertical line, is $\omega = 2\pi\nu = \gamma H$. This is called the *Larmor* frequency. The visualization of rotating magnetic moments is useful when considering the effects of changing magnetic fields on the nuclear spins.

We first discuss NMR in some detail because it is extensively used in biology. Although ESR has been used to obtain important information about biological systems, it is less extensively used and will receive relatively brief consideration.

13.2 NMR SPECTROMETERS

The first NMR spectrometers placed a sample in a fixed magnetic field and applied a fixed radio frequency by means of a coil perpendicular to the field direction. The magnetic field was varied by a coil until resonance was achieved, with the absorption being detected by a third coil. This is analogous to the methods used in visible and ultraviolet spectroscopy. Although the magnetic field applied to the sample is uniform, the actual magnetic field at the nucleus is dependent on a number of environmental factors so that the field must be scanned to find the energy absorption (resonance) condition. We will return to this matter a bit later. A typical spectrum would display the absorption of energy versus frequency, as shown schematically in Figure 13-2.

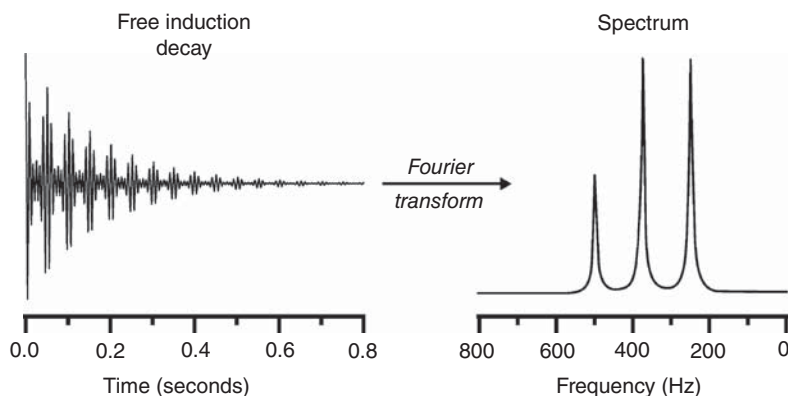


FIGURE 13-2. The free induction decay of the signal from nuclei in a magnetic field is shown (*left*) after a short pulse of radiofrequency radiation is applied to the sample. A Fourier transform of the free induction decay gives the familiar NMR spectrum on the right where the absorption is plotted versus the frequency. Copyright by Professor T. G. Oas, Duke University. Reproduced with permission.

The absorption of radiation is difficult to detect because the populations of the energy states are very similar, as discussed earlier. In practice, this means that relatively high concentrations of the species being observed must be present. The sensitivity of detection depends on the characteristics of the nucleus being observed and its natural abundance. The sensitivity is also enhanced by increasing the energy difference between the two states, which is the reason that instruments with larger magnetic fields are being developed continuously. The proton, ^1H , provides the best sensitivity, and its natural abundance is 99.98%. Consequently, the most extensive measurements have been carried out with proton NMR. Although the natural abundance of ^{31}P is essentially 100%, the sensitivity of detection relative to protons is only about 6%. Probably the second most studied atom with NMR is ^{13}C . Its natural abundance is only about 1%, but its prevalence in biological compounds is very extensive. Furthermore, its abundance in molecules of interest can be enhanced through synthesis or bacterial growth with ^{13}C -enriched compounds. Regrettably, the most abundant isotope of carbon, ^{12}C , does not have a nuclear spin. A variety of other isotopes have been studied with NMR. Most notable for biologists are ^{14}N , ^{15}N , and ^{19}F , which can often be substituted for H in substances of biological interest.

As discussed in Chapter 7, the frequency dependence of the absorption can be transformed into a time dependence and vice versa through Fourier transforms. The most common method of obtaining an NMR spectrum today is to apply a timed radio frequency pulse and then watch the nuclei return to their equilibrium configurations. The time dependence of the return to equilibrium or *free induction decay* (FID) can be transformed from a time-dependent signal into a frequency spectrum, as shown schematically in Figure 13-2. Multiple FIDs can be combined to produce an average FID that has less noise than a single FID. Consequently, Fourier transform instruments are considerably more sensitive than continuous wave instruments so that lower concentrations can be used. The timing and nature of the pulses can be quite complex, but this does not alter the basic concept underlying Fourier transform methods.

13.3 CHEMICAL SHIFTS

Thus far, we have discussed NMR as though the nuclei were isolated in the magnetic field, but the utility of NMR derives from the interaction of the nuclei with surrounding electrons and other nuclei in the molecule. The external magnetic field interacts with the electrons to induce magnetic moments in the electrons that usually oppose the external field. Consequently, the magnetic field at the nucleus is usually lower than the external field. This shielding effect of the electrons can be incorporated into the standard equations by noting that the field at the nucleus, H , in a static magnetic field, H_0 , is

$$H = H_0(1 - \sigma) \quad (13-4)$$

where σ is the shielding constant, typically about 10^{-5} . The value of σ can be positive or negative, depending on whether the magnetic field from the electrons aligns against or with the external magnetic field.

The shielding effect is directly proportional to the strength of the external field, but the same relative change in resonance frequency is observed, regardless of the external field strength. This shielding effect, therefore, can be expressed as a relative change in frequency with respect to a standard, thereby rendering it independent of the external field. These frequency changes are called *chemical shifts*, δ , and are given in units of parts per million (ppm). The chemical shift can be written as

$$\delta = \frac{\nu - \nu_{\text{ref}}}{\nu_{\text{ref}}} 10^6 = (\sigma_{\text{ref}} - \sigma) 10^6 \quad (13-5)$$

where ν is the frequency of the nucleus, ν_{ref} is the frequency of a standard compound, and σ_{ref} is the shielding constant of the standard compound. The most common reference compound for protons is tetramethylsilane, but it is insoluble in water so trimethylsilylpropionate- d_4 is usually used in aqueous media (cf. Ref. 1 for a discussion of chemical shift references). In principle, the chemical shift depends on the

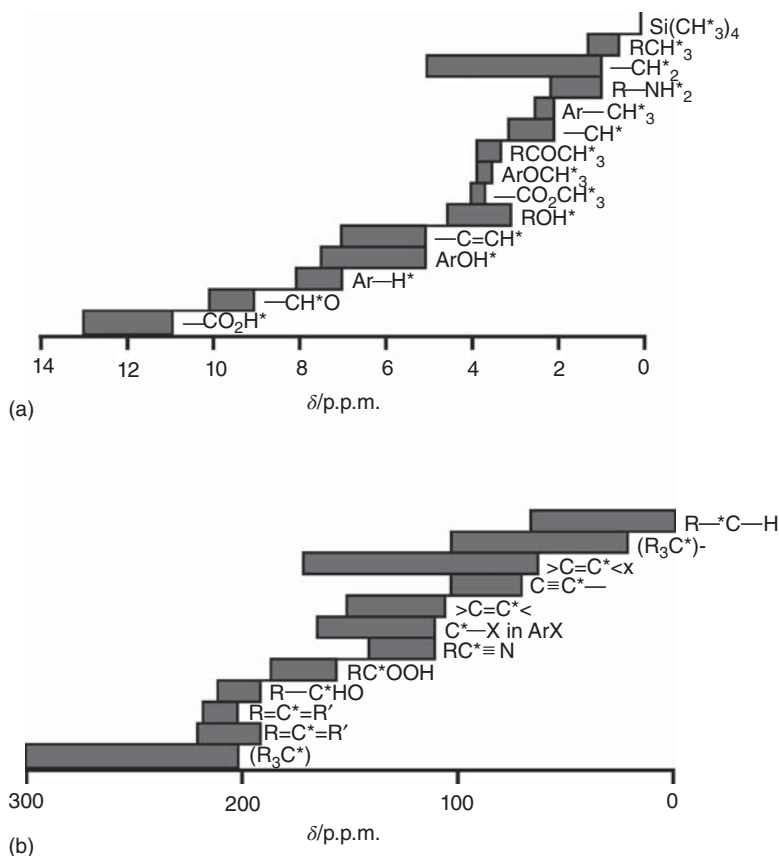


FIGURE 13-3. Range of typical chemical shifts for ^1H and ^{13}C resonances. P. W. Atkins, *Physical Chemistry*, 3rd edition, W. H. Freeman, New York, NY, 1986, p. 489. © 1978, 1982, 1986 by Peter W. Atkins. Used with permission of Freeman.

orientation of the sample with respect to the magnetic field. In liquids, this is generally not a problem because molecules are rapidly tumbling and sampling all possible orientations. For large molecules that tumble relatively slowly, however, the resonances can become so broad that the spectrum is obscured, and special techniques such as TROSY, which is discussed below, are used to interpret the spectra. In solids, special conditions are required to obtain high-resolution spectra.

The electron density at the nucleus is often the dominant factor in determining the chemical shift. A high electron density creates a large shielding, and the applied magnetic field must be increased to get resonance: this results in an upfield shift and a decrease in the magnitude of δ because reference compounds are generally highly shielded. Conversely, a low electron density at the nucleus causes a downfield shift and increase in δ . The range of chemical shifts for ^1H and ^{13}C in various compounds is shown in Figure 13-3.

In principle, the chemical shifts of nuclei in proteins should provide information about the protein structure. However, chemical shifts alone are not sufficient to determine protein structure. The average chemical shifts of various nuclei for amino acids in denatured proteins are given in Table 13-2. Note that they are quite similar for all of the amino acids. When secondary structures such as α -helices or β -sheets are present, changes in chemical shifts occur so that the amounts of various secondary structures present can be inferred from the NMR spectra.

TABLE 13-2. Average Chemical Shifts of Random Coil Amino Acids (ppm)

Amino acid	α - ^1H	Amide- ^1H	α - ^{13}C	Carbonyl- ^{13}C	Amide- ^{15}N
Ala	4.33	8.15	52.2	177.6	122.5
Cys	4.54	8.23	56.8	174.6	118.0
Asp	4.71	8.37	53.9	176.8	120.6
Glu	4.33	8.36	56.3	176.6	121.3
Phe	4.63	8.30	57.9	175.9	120.9
Gly	3.96	8.29	45.0	173.6	108.9
His	4.60	8.28	55.5	174.9	119.1
Ile	4.17	8.21	61.2	176.5	123.2
Lys	4.33	8.25	56.4	176.5	121.5
Leu	4.32	8.23	55.0	176.9	121.8
Met	4.48	8.29	55.2	176.3	120.5
Asn	4.74	8.38	52.7	175.6	119.5
Pro	4.42	—	63.0	176.0	128.1
Gln	4.33	8.27	56.0	175.6	120.3
Arg	4.35	8.27	56.0	176.6	120.8
Ser	4.47	8.31	58.1	174.4	116.7
Thr	4.35	8.24	62.0	174.8	114.2
Val	4.12	6.19	62.2	176.0	121.1
Trp	4.66	8.18	57.6	173.6	120.5
Tyr	4.55	8.28	58.0	175.9	122.0

Reproduced with permission from D. S. Wishart and B. D. Sykes, Chemical shifts as a tool for structure determination, *Methods Enzymol.* **239**, 363 (1994). © 1994, with permission of Elsevier.

Some of the largest changes in chemical shifts in proteins and nucleic acids are observed for aromatic rings of nucleotides, tyrosine, phenylalanine, and tryptophan. These shifts are due to the interaction of the external magnetic field with the delocalized electrons of the aromatic ring. Nuclei above or below the ring usually have decreased chemical shifts, whereas those near the edges have increased chemical shifts. These effects are called *ring currents* and can cause unusually large chemical shifts. These chemical shifts can provide information about the structure of macromolecules and about alterations in structure due to changes in the environment such as temperature and the addition of various chemical agents.

13.4 SPIN-SPIN SPLITTING

The chemical shift is caused by interactions between the nucleus and nearby electrons. A conceptually different interaction is transmitted between nearby nuclei by intervening electrons participating in chemical bonds. Basically, the spin state of a neighboring nucleus alters the shielding a nucleus experiences. This effect is smaller than typical chemical shifts and is called *spin-spin splitting* and is often referred to as scalar coupling. Unlike chemical shifts, the magnitude of spin-spin coupling is independent of the magnitude of the applied magnetic field.

This effect is most easily understood by considering the proton NMR spectra of ethanol shown in Figure 13-4. The low-resolution spectrum (left) shows the three peaks that might be expected on the basis of our discussion of chemical shifts, namely, CH_3 protons, CH_2 protons, and the OH proton. The ratio of the areas under the resonances is proportional to the number of protons in each chemical environment, 3:2:1. The high-resolution spectrum (right) shows that these three resonances are multiplets of peaks due to spin-spin splitting. The number of peaks within a multiplet is determined by the number of spin orientations of neighboring nuclei. In this case, the spins of the two hydrogens on the methylene (CH_2) carbon can have four possible arrangements of orientations

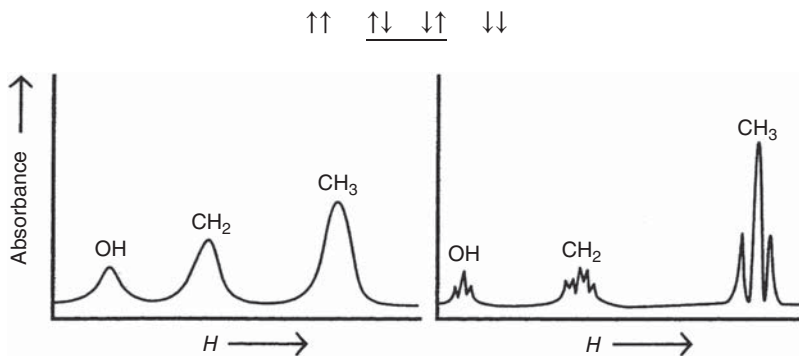


FIGURE 13-4. Schematic representation of the NMR spectrum of dry ethanol at low (*left*) and high (*right*) resolution. The ratio of the areas under the peaks is 3:2:1. The octet expected for the methylene protons (*right*) is not shown as the splitting of the four peaks is very small.

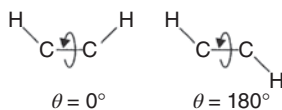
The middle two arrangements are equivalent so that the methyl proton resonances are split by the neighboring CH_2 into three resonances with relative areas under the peak of 1:2:1. The OH proton, also next to the methylene carbon, is split into three peaks. Similarly, the methyl proton can have three different orientations, with the two underlined groups being equivalent:



The methylene resonance, therefore, is split into four peaks, with area ratios of 1:3:3:1. The hydroxyl proton will split these four peaks further into an octet, but this detail is difficult to see in the spectrum. In order for scalar coupling to occur, the two nuclei must have distinct resonance peaks. Thus, the methyl protons will not split each other's resonances.

Spin-spin splitting is of special importance in considering the NMR spectra of proteins because coupling between the α -C proton and the N proton of the amide bond occurs. Scalar coupling can also occur between different nuclei, for example, a proton and ^{15}N . The magnitude of the spin-spin splitting is called the coupling constant and is designated by J . The number of covalent bonds separating the nuclei in question is often appended as a prior superscript, and a post subscript designates the atoms involved. For example, the coupling constant for the amide proton and the α -C proton would be written as $^3J_{\text{HN-H}\alpha}$.

The spin-spin coupling constant for protons on adjacent atoms varies considerably, from about 0 to 10 Hz. This variation is due to different torsional angles between the protons as defined below:



If the torsional angle is 90° , the coupling is close to zero, whereas when it is 0° or 180° , it is about 10 Hz. If free rotation about the bond occurs, an average coupling constant is obtained. The dependence of the coupling constant on the torsional angle is given by the Karplus equation:

$$J = A + B \cos \theta + C \cos^2 \theta \quad (13-6)$$

The constants A , B , and C can be calculated or established from measurements with molecules having known dihedral angles. If free rotation about the bond occurs, the coupling constant will be some average of the possible dihedral angles. However, if the bond is constrained, such as in a peptide linkage that has double bond character or in a folded macromolecular structure, the calculated dihedral angle can provide useful information about the structure. This is especially true when used in conjunction with other information.

This simple picture of spin-spin splitting is not rigorous. It provides a useful and adequate explanation for relatively simple situations but breaks down when considering multidimensional NMR, which is discussed later.

13.5 RELAXATION TIMES

Thus far, we have not dealt explicitly with the time scale for nuclear spin transitions. If we think of a pulsed NMR experiment, we can envisage nuclear spins being oriented in a specific direction by the pulse, with the precession of spins about the direction of the field. When the pulse is turned off, the nuclei will come to equilibrium with regard to their environment. The rate at which a particular nuclear spin returns to equilibrium depends on interactions with other nuclear spins, and this in turn will depend on the fluctuating fields experienced as the molecules tumble. This mode of relaxation involves a change of energy between the spin systems and their environment. The return of a spin population in a magnetic field to its equilibrium population follows first-order kinetics, and the reciprocal of the first-order rate constant is called the *spin-lattice relaxation time*, T_1 . Typically, it is in the range of tenths of a second to seconds for protons.

Simply put, the value of T_1 depends on the interactions of a nuclear spin with its neighbors, including the solvent, and how fast the molecule rotates. The molecular rotation is characterized by a rotational correlation time that is basically a measure of the rotational diffusion constant of a molecule. Measurements of T_1 can, in fact, be used to determine rotational diffusion constants of rigid molecules. In some cases, chemical exchange of nuclei in different environments can contribute to T_1 , although this is somewhat unusual. In general, anything that gives rise to magnetic fluctuations in the environment can contribute to this relaxation, for example, unpaired electrons or dissolved oxygen.

A second mode of nuclear spin relaxation is possible that does not involve the exchange of energy of the magnetic moment with its environment. In terms of the picture of a nuclear spin being oriented in a field and precessing about the field direction, the spin-lattice relaxation time can be thought of as characterizing the return of the spin orientations to their equilibrium positions. The *spin-spin relaxation time*, T_2 , on the other hand, is associated with the rate of precession about the field direction. Basically, T_2 is a measure of alterations in the precession frequency during nuclear spin relaxation. This alteration is different for different nuclear spins so that the rates of precession change with respect to each other. This is called a loss of phase coherence. This can be viewed as an exchange of energy within the spin system. It does not change the net population of the excited states.

The dominant factor determining T_2 is the rate of molecular tumbling. The molecular tumbling effect is quantitatively different for T_2 and T_1 . Because T_2 is generally much shorter than T_1 in liquids, the line widths of spectra are determined by T_2 , $1/T_2 = \pi\nu_{1/2}$ where $\nu_{1/2}$ is the peak width at one-half of its maximum value. As previously stated, rapidly tumbling molecules produce relatively sharp lines whereas slowly tumbling molecules have relatively broad lines.

Chemical reactions provide one of the most interesting examples of how T_2 can be altered. Consider the simple chemical reaction



Assume that a specific proton has a different chemical shift in A and B. (This means they have a different Larmor frequency.) If the chemical exchange rate is very slow compared to the difference in chemical shifts (i.e., smaller than the difference in Larmor frequencies), two distinct peaks will be seen in the NMR spectrum, as shown in Figure 13-5 (top). At the other extreme, if the chemical exchange rate is very fast relative to the difference in the chemical shift frequencies, states A and B will interconvert many times during the NMR experiment, and the two frequencies are effectively averaged. Consequently, only a sharp single line will be seen, located at the average of the two lines weighted by the relative populations of A and B (Fig. 13-5, bottom). In the intermediate cases, where the rate of the chemical reaction is comparable to the difference in chemical shift frequencies, the two lines will broaden, coalesce into a single broad peak, and finally sharpen, as shown in Figure 13-5. For a single line, the spin-spin relaxation time can be written as

$$\frac{1}{T_2} = \frac{1}{T_{2A}} + \frac{1}{\tau_A} \quad (13-8)$$

where T_{2A} is the relaxation time of the A state without significant chemical exchange and τ_A is the relaxation time for the chemical reaction. We will not delve into

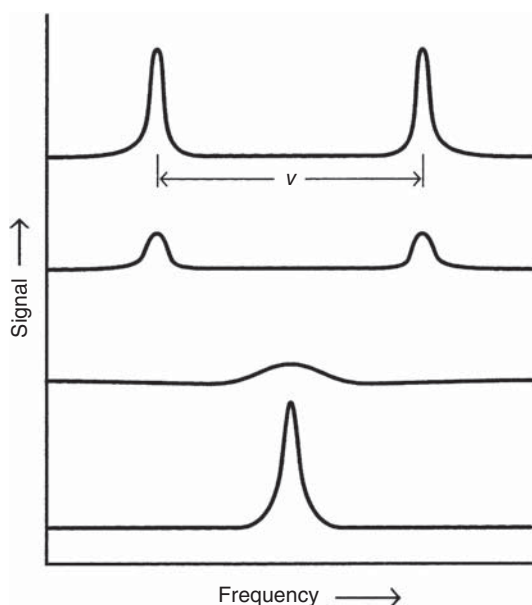


FIGURE 13-5. Schematic drawing of the effect of chemical exchange on NMR spectra. In the top spectrum, resonances from two protons are shown. In this case, the rate of exchange between the two environments is much less than ν , or $1/\tau_A \ll \nu$, whereas in the lowest spectrum, $1/\tau_A \gg \nu$, so that only a single resonance is seen that is the average of the positions in the upper spectrum. The two spectra in the middle represent the cases where $1/\tau_A \sim \nu$, with $1/\tau_A$ becoming progressively larger from the top spectrum to the bottom spectrum.

the details, but the line shape can be quantitatively analyzed to determine the rate constants for the chemical reaction. The time scale of the reactions that can be studied is determined by the chemical shift difference between the two states. For example, if the difference is 100 Hz, then rates of chemical reactions in the range of $1/100 = 10^{-2}$ s can be studied. Reactions such as the rate of exchange of hydrogens between proteins or amino acids and water have been investigated with this method.

13.6 MULTIDIMENSIONAL NMR

Thus far, we have considered what is now called one-dimensional NMR, namely determining the spectrum for a specific nucleus by scanning the magnetic field or by analysis of the frequencies associated with the free induction decay following a frequency pulse. One-dimensional analysis has been invaluable in determining the structures of small molecules and can also be used to obtain information about macromolecules. However, in order to get definitive structural information about macromolecules, multidimensional NMR is necessary. The genesis of this field was in the early 1970s, and its vigorous evolution continues to this day. The underlying principle of multidimensional NMR is to find “cross peaks” that link two resonances. This linkage can be either through space or through a small number of chemical bonds. Finding these cross peaks allows the spatial relationships to be determined between the nuclei responsible for the two resonances. These connections between resonances are sometimes called coherence pathways.

A detailed presentation of multinuclear NMR is beyond the scope of this text. However, the concepts can be understood by considering two-dimensional NMR in a qualitative manner. Two-dimensional NMR can be viewed as the assembly of one-dimensional spectra in an array. The experiments consist of four stages, illustrated in Figure 13-6. In the preparation phase, a frequency pulse is applied to the system. An evolution phase of length t_1 then occurs and is followed by a mixing period ending in a second pulse. (A pulse may also be applied between the evolution and mixing phases, depending on the specific two-dimensional experiment being carried out.) Finally, a free induction decay occurs for a time t_2 in the data acquisition

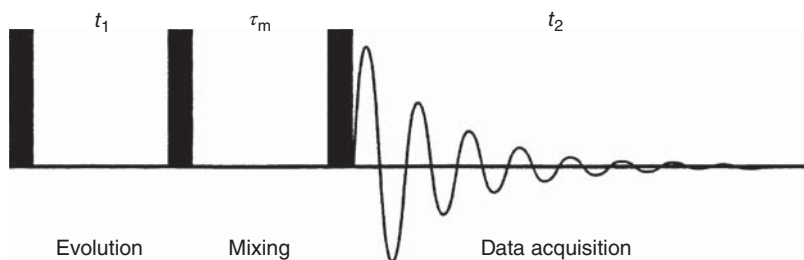


FIGURE 13-6. Schematic representation of a two-dimensional NMR experiment in which a radiofrequency pulse (black bar) is applied to the sample initially. After a period t_1 , a second pulse is applied, followed by a mixing time, τ_m . The data are acquired after a final pulse. The number of pulses and the various times depend on the type of experiment being carried out.

phase. In the actual experiment, t_1 is varied incrementally and the signal is collected for the period t_2 , ultimately giving a large number of time points. This experiment is repeated many times and signal averaged. The nature of the pulses will depend on the specific experiment. Qualitatively, the pulses flip the spins (magnetic moments) in the field that is created, and the end of the mixing phase flips them again. In the evolution phase, the magnetic moments associated with the nuclear spin will partially return to equilibrium at a rate that depends on T_1 and T_2 . The return to equilibrium also occurs in the collection phase. By contrast, one-dimensional NMR uses a single pulse, followed by free induction decay and data collection.

The design of these experiments can be quite tricky, as it depends on what connections between spin states are being probed and what the relaxation times are. The representation of the result is done by carrying out a Fourier transform on both the t_1 and t_2 data sets and converting them into a two-dimensional plot of ν_1 versus ν_2 with the third dimension being the amplitude of any resonance peaks that are observed. A very simple illustration is given in Figure 13-7 for two interacting spin systems. Peaks that occur on the diagonal have the same frequency in both dimensions and correspond to the one-dimensional spectrum, whereas those

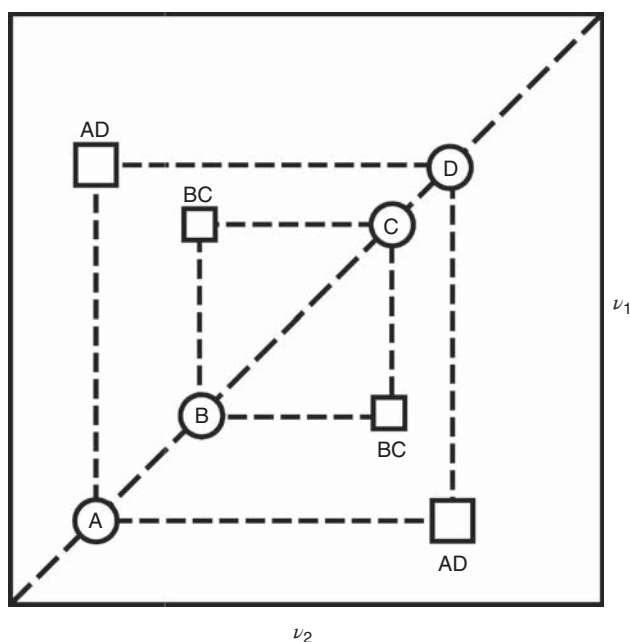


FIGURE 13-7. Schematic representation of a two-dimensional NMR spectrum for two spins. The frequency axes labeled ν_1 and ν_2 are the positions of the resonances for the two different spins. The diagonal (circles) corresponds to the one-dimensional spectrum. If coherence transfer occurs during the mixing period, off-diagonal resonances (squares) will be seen. In this example, coherence transfer occurs between nuclei with resonances at A and D and between nuclei with resonances at B and C.

occurring off the diagonal represent cases where different spin systems interact during the mixing period (coherence transfer). Contours are usually used to indicate the amplitudes of the resonance peaks, rather than a third dimension. The interactions (coherence transfer) between nuclear spins can be either homonuclear (same nuclei) or heteronuclear (different nuclei).

One of the first two-dimensional experiments carried out was COSY (CORrelated Spectroscopy). This experiment identifies pairs of nuclei that are linked by scalar coupling (spin–spin splitting connectivities). The interaction between these nuclei occurs during the relaxation taking place in the mixing phase and results in cross peaks in the spectrum. Scalar coupling (off-diagonal peaks) between nuclei means that they are within three covalent bonds. A typical COSY spectrum is shown in Figure 13-8 for λ Cro repressor, a DNA-binding protein that regulates phage development.

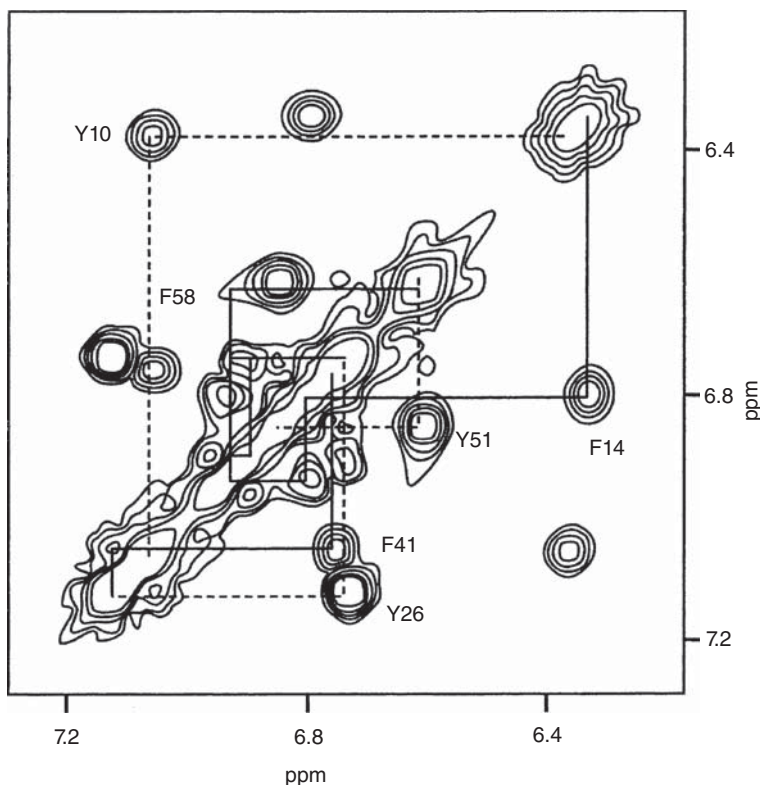


FIGURE 13-8. Example of a COSY spectrum for the protons of the aromatic spin systems in λ Cro repressor protein. Tyrosine rings are connected with a dashed line and phenylalanine rings by a solid line. The diagonal and cross peaks can be easily seen. Reprinted in part with permission from P. L. Weber, D. F. Wemmer, and B. R. Reid, *Biochemistry* **24**, 4553 (1985). © 1985 by American Chemical Society.

One of the most important two-dimensional NMR experiments is Nuclear Overhauser Effect Spectroscopy (NOESY). This is a through space interaction that takes place because of the interactions between the magnetic dipoles of two nuclear spins. In terms of our previous discussion, this primarily involves the T_1 mode of relaxation and is coupled with rotational motion. The NOE is the NMR equivalent of fluorescence resonance energy transfer discussed in Chapter 10. In fact, an NOE can be observed in one dimension. If a sufficiently large magnetic field is applied at the resonance condition of a given nucleus, the spin system becomes saturated, that is, the ground and excited states become equally populated so that no more energy can be absorbed. If the magnetic moment of this nucleus is sufficiently close in space to another magnetic moment, energy can be transferred. This perturbs the intensity of the resonance of the nucleus to which energy is transferred and decreases the resonance intensity of the nucleus that was initially irradiated. The interaction between magnetic dipoles varies as the inverse sixth power of the distance between the two dipoles so that only nuclear spin systems that are very close to each other give rise to NOEs. In practice, this means distances of 5 Å or less.

The NOESY experiment is quite similar to the two-dimensional experiment described previously, except that a pulse is applied at both the beginning and the end of the mixing period, with the time of the mixing period being constant. An example of a NOESY spectrum is given in Figure 13-9 where proton–proton NOEs for a complex of DNA and an antibiotic, distamycin A, are presented. Because of the distance dependence of this effect, the distance between spins can be estimated from the cross peaks. However, because only the distance, not the direction, is derived from these measurements, many distances must be determined to arrive at a unique structure.

Another NMR technique, residual dipolar coupling (RDC), is now commonly used to provide information on the relative orientations of the nuclei and the directions of bonds connecting them in macromolecules. In high-resolution NMR, the interactions between the magnetic dipoles of nuclei are averaged to zero by the rapid tumbling of molecules, which produces narrow spectral resonances. However, additional structural information can be obtained by partially orienting samples in the magnetic field. In this case, the coupling between nuclear magnetic dipoles can be measured as RDC. These measurements give information about the relative orientations of the nuclei. Moreover, these interactions extend over a longer distance than NOE effects. Knowledge of the relative orientations of nuclei can be translated directly into structural information. Thus comprehensive RDC measurements can be used to determine macromolecular structures. They can also be used to obtain dynamic information, a topic that will not be pursued here.

One of the most useful types of multidimensional spectra is HSQC (Heteronuclear Single Quantum Correlation), an example of which is shown in Figure 13-10 for a protein involved in proton transport across a membrane (subunit c of ATP synthase from *E. coli*). This spectrum uses both ^1H and ^{15}N and selectively detects only pairs of covalently attached nuclei. Each spot in the contour map represents such a pair, with the position on the horizontal axis representing the proton resonance frequency and the position on the vertical axis representing the resonance frequency of

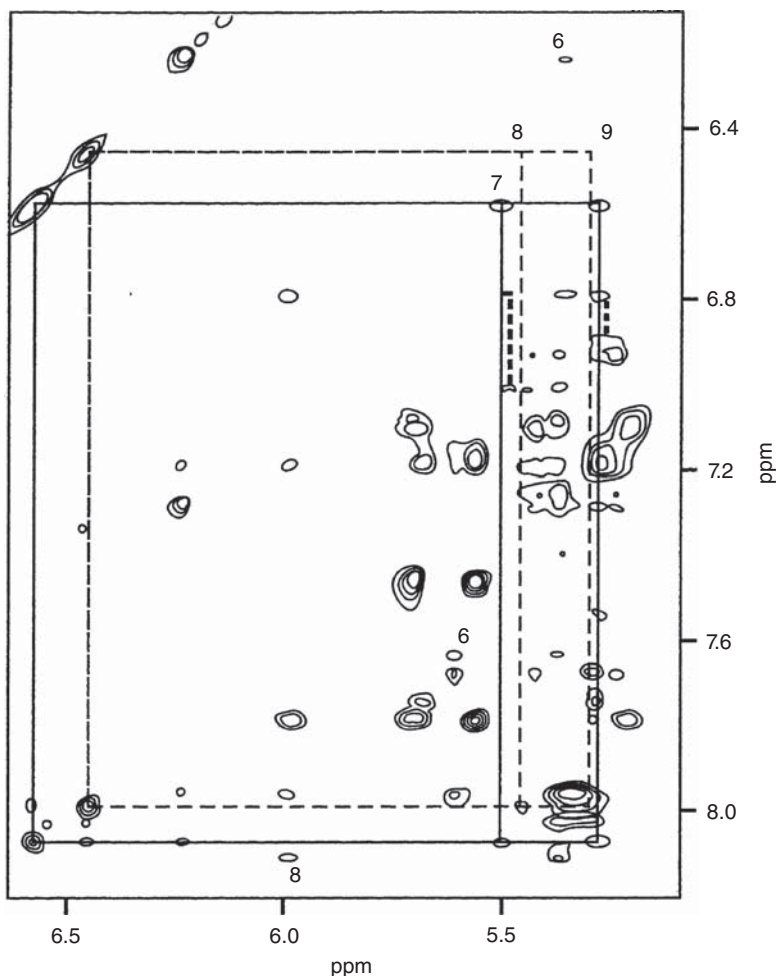


FIGURE 13-9. Example of a NOESY spectrum for a complex of distamycin A and DNA. Aromatic C6H resonances of adenine and guanine and C2H resonances of adenine are shown along the vertical axis. The C1'H resonances are shown along the horizontal axis. Sets of sequential connectivities are denoted by dotted, dashed, and solid lines. Reprinted in part with permission from J. G. Pelton and D. F. Wemmer, *Biochemistry* **27**, 8088 (1988). © 1988 by American Chemical Society.

the nitrogen nucleus. Because every amino acid, except for proline, has a backbone amide, essentially every residue is represented in this spectrum. Side chains containing amides will give rise to additional resonances. This gives direct information about which nitrogen is coupled to which hydrogen. Furthermore, the resonances are usually quite well resolved. The HSQC spectrum provides information that is useful for assigning the observed resonances to specific amino acid residues.

Extending NMR to dimensions greater than two involves similar concepts. Multiple pulses and even field gradients are used. These dimensions can be in terms of

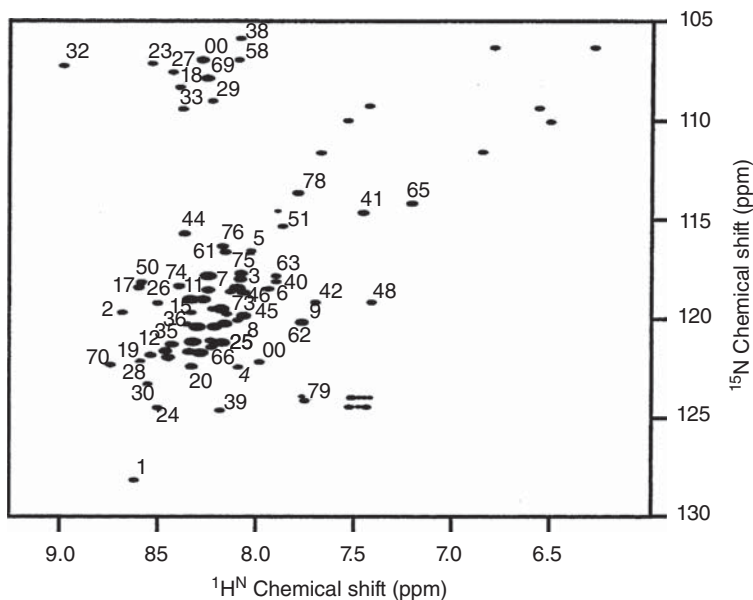


FIGURE 13-10. Example of an HSQC spectrum. All 76 backbone amide cross peaks are shown for the c subunit protein of ATP synthase from *E. coli*. This membrane-bound protein is involved in proton transport across the membrane. Reprinted in part from M. F. Girvin, V. K. Rastogi, F. Abildgaard, J. L. Markley, and R. H. Fillingame, *Biochemistry* **37**, 8817 (1998). © 1998 by American Chemical Society.

different nuclei and/or combining two-dimensional experiments. The most common nuclei studied in addition to ^1H are ^{13}C and ^{15}N . In these cases, proteins and nucleic acids enriched in the nucleus of interest can be used to enhance the sensitivity.

Transforming the NMR results for a macromolecule into a structure is straightforward but not easy. First, the resonances have to be assigned to specific nuclei within the structure. This can be done by analysis of scalar coupling and sequential NOEs. The three-dimensional structure is derived from the distances determined from NOE experiments, RDC measurements, and dihedral angle information determined from spin-spin splitting and the Karplus equation (Eq. (13-6)). In practice, the information gleaned from the NMR spectra provides hundreds of constraints on the structure. A number of computer programs have been written that take the constraints and convert them into a family of structures, usually very similar, consistent with the constraints. These programs involve sophisticated theory as well as data analysis to arrive at final structures. NMR spectroscopy is a very powerful tool for determining protein structures in their biologically active conformations.

As proteins increase in size, their rate of rotation slows down, and NMR spectra broaden. This limits the size of proteins for which structures can be determined. However, several recent advances have greatly enhanced the use of NMR for structure determination. Resolution can be greatly improved by the use of *transverse relaxation optimized spectroscopy* (TROSY). This technique, which requires a relatively

high magnetic field, selects a narrow component of a broad line by cancelling out various components of the broad line, thereby producing a narrow line. In addition, the sensitivity of NMR has been greatly enhanced by the use of cryogenic probes cooled to near liquid helium temperatures. This allows the study of more dilute solutions of macromolecules. In fact, it is now possible to determine NMR spectra in live cells, permitting the direct study of physiological systems (2).

The size of macromolecules whose structure can be determined with NMR increases year by year. At the present time, structures can be determined for molecules as large as 50,000–75,000 molecular weight, and features of molecules as large as 200,000 molecular weight can be discerned.

13.7 MAGNETIC RESONANCE IMAGING

One of the most remarkable advances in diagnostic medicine has been the evolution of magnetic resonance imaging (MRI). Although X-rays readily distinguish hard objects such as bones, they do not distinguish soft tissue structure. MRI, on the other hand, provides excellent images of tissues and is able to distinguish between various types of tissues. Protons in water are the primary nucleus used for detection, although applications with ^{13}C , ^{31}P , and ^{19}F have been developed. The principle underlying MRI is the use of a magnetic field gradient. Since the resonance frequency is directly proportional to the magnetic field, the frequency of the resonance will depend on its location in the magnetic field. The intensity of the absorption is dependent on how many protons are present. If a linear magnetic field gradient is applied, the position of the resonance will change, also linearly, as the field is varied. Thus, a plot of the amplitude of the resonance versus frequency is equivalent to a plot of the integrated number of protons versus distance. A series of cross sections can be obtained by rotating the sample in the field, or by moving the field around the sample. These cross sections can then be reconstructed to give a three-dimensional image.

In soft tissue, the amount of water varies for different tissues, so the density of protons varies. In addition, the various tissues are characterized by significantly different T_1 values. The difference in proton density can be shown in reconstructions by varying the darkness of the shading. This can be seen in Figure 13-11, where the MRI image of an adult human brain is shown. In addition to being able to distinguish various soft tissues, MRI is noninvasive, as contrasted to X-rays or injections of foreign substances, including radioactive isotopes, required for other imaging techniques. Mention also should be made of functional MRI. In these studies, metabolism and neurological activity can be directly monitored in living subjects.

13.8 ELECTRON SPIN RESONANCE

We now return briefly to ESR. Unlike NMR, ESR is not observed for most materials. This is because electrons are usually paired and consequently have no net magnetic moment. However, free radicals and other paramagnetic substances have unpaired

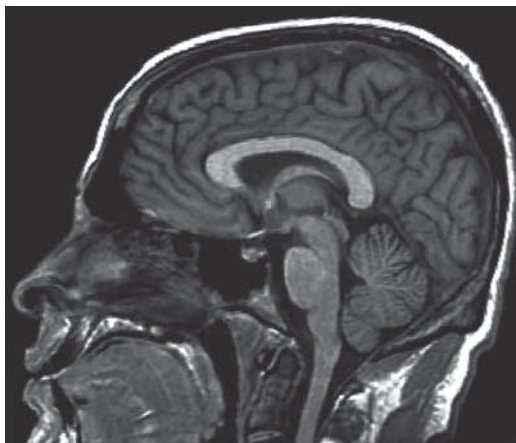


FIGURE 13-11. MRI of an adult human brain. Courtesy of Professor Scott Huettel, Duke University Brain Imaging and Analysis Center. Reproduced with permission.

electrons that give rise to ESR spectra. Because the frequencies associated with transitions between the energy levels of unpaired electrons in a magnetic field are in the microwave region, special techniques are required for placement of the sample in the magnetic field. Conceptually, the experiment is the same as for NMR. The magnetic field is varied until resonance is found.

The usefulness of ESR in biological systems arises from the interactions between nuclear spins and the electron spin. This gives rise to “hyperfine structure” in the spectrum. For example, if a neighboring nucleus has a nuclear spin of $1/2$, it will have two orientations in the field and will split each of the two energy states of the electron into two, as shown in Figure 13-12. Because of quantum mechanical selection rules, only two transitions between these four energy states are allowed. If the nuclear spin is 1, six energy levels are created and three transitions between the levels are allowed. The hyperfine structure can provide information about the environment of the paramagnetic species.

As with NMR, the sharpness of the spectrum provides information about the rotational mobility of the paramagnetic species. However, the time scale for rotation is much different than NMR because the frequencies are much higher. For ESR, rotational correlation times of approximately 10^{-9} s influence the line width, whereas for NMR the time scale is about 10^{-6} s. For reasons we will not dwell on here, the spectra are usually presented as the derivative of the amplitude versus the field (Fig. 13-13).

One of the important developments in the application of ESR to biological systems was the synthesis of stable free radicals that could be reacted with macromolecules and membranes, both noncovalently and covalently (3). The most common element of these spin labels is the nitroxide free radical (Fig. 13-14). Because the ^{14}N nucleus, with a nuclear spin of 1, is next to the free radical, three bands are seen in the spectrum. The derivative of the resonances is three up and down peaks. This is illustrated in Figure 13-14 where ESR spectra are shown for nitroxides under various conditions of rotational mobility. The effect of molecular motion on the spectra of the free radical

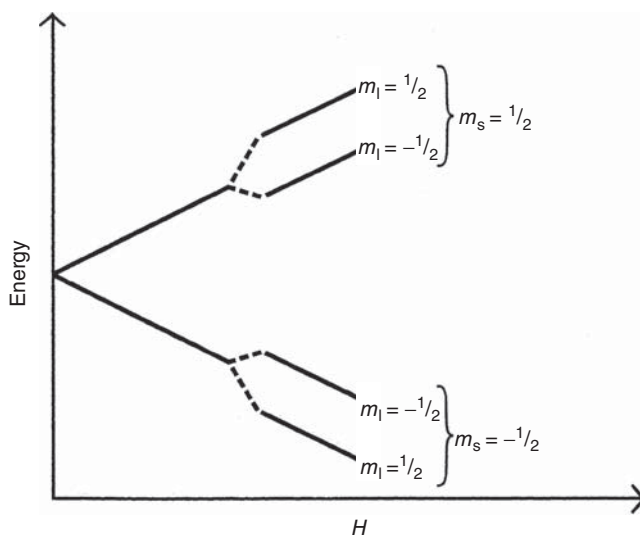


FIGURE 13-12. Hyperfine splitting of the energy levels of an electron in a magnetic field, H , by a nuclear spin. The electron has a spin quantum number, m_s , of $\pm 1/2$ and the nucleus has a spin quantum number, m_l , of $\pm 1/2$.

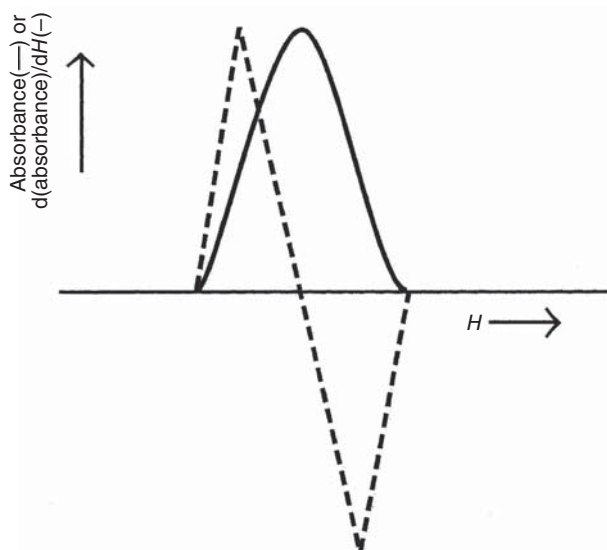


FIGURE 13-13. Schematic representation of a derivative spectrum. An absorbance peak is shown as a solid line and its derivative is shown as a dashed line. ESR spectra are usually presented as derivatives of the absorption.

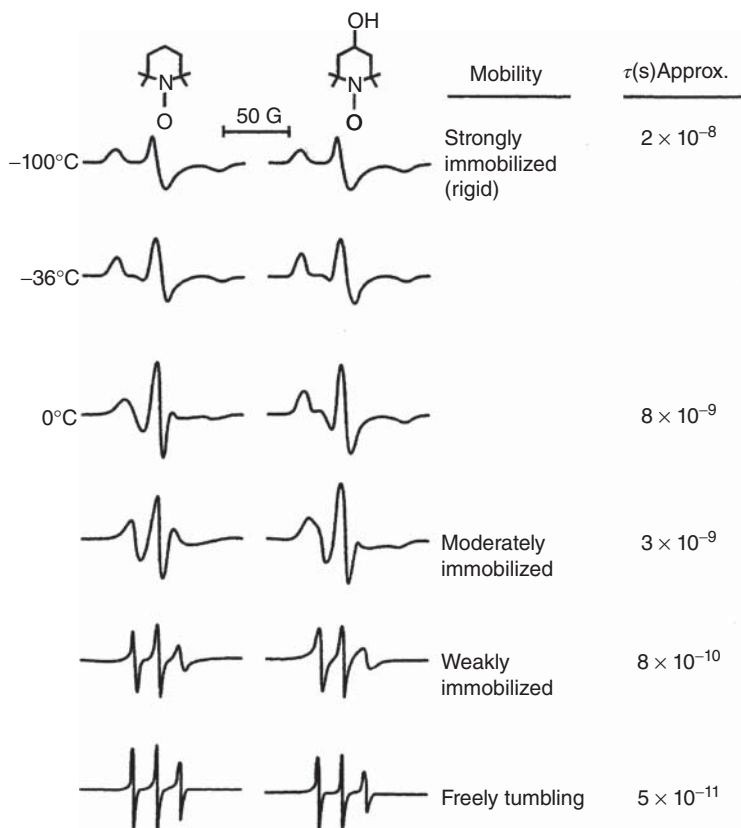


FIGURE 13-14. ESR spectra of spin labels whose structures are shown at the top of the figure. The effect of viscosity on the line shapes and rotational correlation times, τ , is shown. Reproduced with permission from R. A. Dwek, *Nuclear Magnetic Resonance in Biochemistry*, Clarendon Press, Oxford, England, 1973, p. 289. Adapted from data in P. Jost, A. S. Waggoner, and O. H. Griffith, Spin Labeling and Membrane Structure, in *Structure and Function of Biological Membranes* (L. Rothfield, ed.), Academic Press, New York, 1971, p. 83. 1971, with permission of Elsevier.

is clearly illustrated. As these spectra illustrate, spin labels can provide information about the rotational mobility of the macromolecule to which they are bound. However, it should be noted that the probe mobility can be due to macromolecule rotation, segmental motion of the macromolecule, or simply rotation of the probe in the site to which it is bound. The interpretation of what motion is being observed is not always straightforward.

ESR spectra can also provide information about the polarity of the spin label environment. The extent of splitting of the hyperfine structure depends on the effective dielectric constant of the environment. Thus, for example, the splitting is quite different in a biological membrane than in solution. When bound to a membrane, the

splitting will be different for a spin label close to a polar head group and for a spin label buried in the hydrocarbon chains.

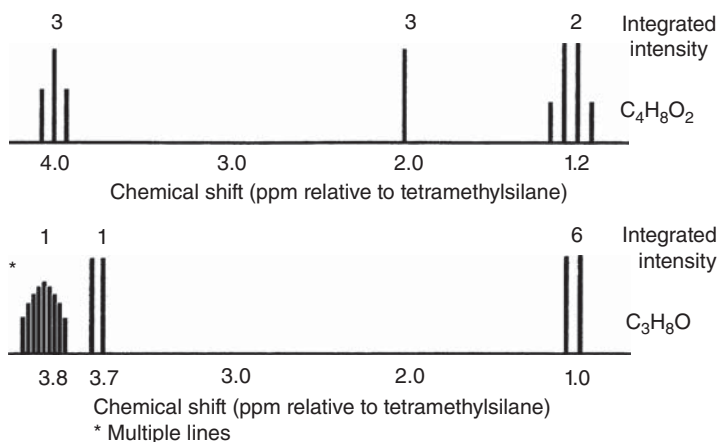
This chapter is a relatively brief introduction to NMR and ESR. Many texts are available that provide more detailed descriptions of NMR (cf. Refs. 4–8) and ESR (cf. Ref. 9). Unfortunately, most of these are not easy reading and require a working knowledge of quantum mechanics. Applications of these techniques to biological systems are presented in the next chapter.

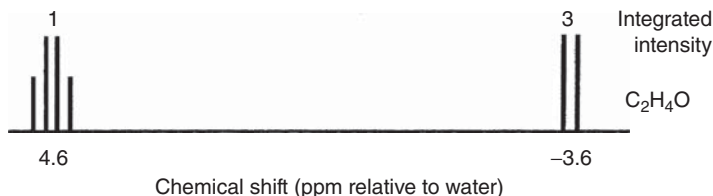
REFERENCES

1. D. S. Wishart and B. D. Sykes, *Methods Enzymol.* 239, 363 (1994).
2. K. E. Robinson, P. N. Reardon, L. D. Spicer, *Methods in Molecular Biology* 831, 261 (2012).
3. S. Ohnishi and H. M. McConnell, *J. Am. Chem. Soc.* 87, 2293 (1965).
4. P. J. Barrett, J. Chen, M. -K. Cho, J. -H. Kim, Z. Lu, S. Mathew, D. Peng, Y. Song, W. D. Van Horn, T. Zhuang, F. D. Sönnichsen, and C. R. Saunders, *Biochemistry* 52, 1303 (2013).
5. I. R. Keckner and M. P. Foster, *Biochim. Biophys. Acta* 1814, 942 (2011).
6. T. L. James and N. J. Oppenheimer, *Methods Enzymol.* 239 (1994).
7. J. Cavanaugh, W. J. Fairbrother, A. G. Palmer, III, and N. J. Skelton, *Protein NMR Spectroscopy: Principles and Practice*, 2nd edition, Academic Press, San Diego, CA, 2008.
8. T. L. James, V. Dötsch, and U. Schmitz eds., *Methods Enzymol.* 238–239 (2001).
9. J. A. Weil and J. R. Bolton, *Electron Paramagnetic Resonance: Elementary Theory and Practical Applications*, Wiley, New York, 2007.

PROBLEMS

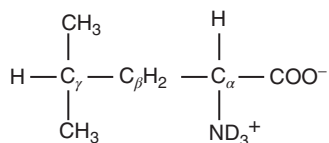
13-1. Deduce the structure of the compounds below from their schematic NMR spectra. Indicate which protons are assigned to each resonance.





- 13-2.** Nucleotide phosphates are important biological molecules. Sketch the ^{31}P (spin 1/2) NMR spectra of a nucleotide monophosphate, diphosphate, and triphosphate. Order the resonances in terms of chemical shift, with zero on the right-hand side. The phosphates are designated as α , β , and γ with α being closest to the sugar and γ furthest away. For the triphosphate, the γ position has the lowest electron density and the β position has the highest.

- 13-3.** The structure of L-leucine in D_2O is

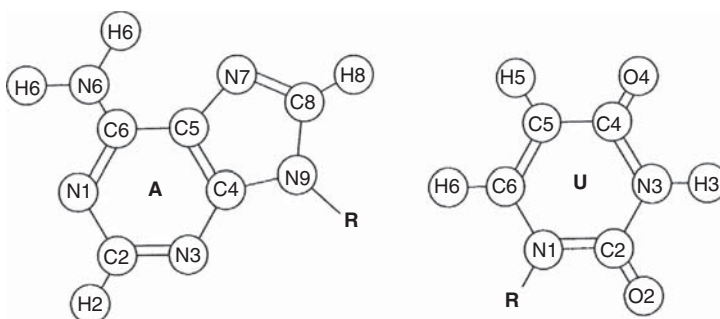


- Sketch the one-dimensional proton NMR spectrum. The approximate chemical shifts are 1.0 for the methyl groups, 1.4 for the γ hydrogen, 1.5 for the β hydrogens, and 3.3 for the α hydrogen. Assume that the methyl hydrogens and β hydrogens are not resolved, that is, all of the hydrogens in each of the two classes have a single resonance peak. Indicate the integrated intensity of the resonances for each resonance.
 - Sketch the proton COESY spectrum of L-leucine, indicating the coupling that should give rise to off-diagonal peaks.
- 13-4.** Two common structures found in proteins are the α -helix and β -sheet (parallel and antiparallel), as discussed in Chapters 3 and 9. The approximate distances in Å between protons for the two structures are given below:

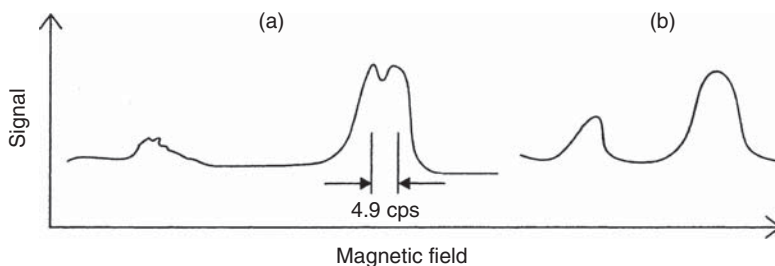
	α -Helix	β -Sheet
Amide H–amide H	2.8	4.2
α -Carbon H–amide H	3.5	2.2
α -Carbon H–(amide H) $_{i+3}$	3.4	>5

The subscript $i + 3$ means the α -carbon is $i + 3$ residues away from the amide proton. Indicate how the measurement of NOEs could be used to distinguish these structures by considering the expected relative intensities of the NOEs for the three classes of protons in the table.

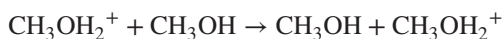
- 13-5.**
- Use Eq. (13-3) to construct a plot of the resonant frequency for an electron as a function of the field strength from 0 to 4 Tesla.
 - Show schematically the splitting expected in the energy levels if the unpaired electron is located next to a nuclear spin of $3/2$.
- 13-6.** An RNA structure containing uracil and adenine was studied by NMR. The structures of these two bases are given below.



- In a $^2J_{NN}$ COSY-type experiment with ^{15}N -enriched RNA, scalar couplings were observed between N3 of uridine and N1 of adenine [A. J. Dingley and S. Grzesiek, *J. Am. Chem. Soc.* **120**, 8293 (1998)]. In addition, in a NOESY-type experiment, an NOE was observed between H3 of uridine and H2 of adenine. Postulate a structure for the hydrogen-bonded dimer formed between uridine and adenine.
 - As the pH of the solution is lowered, the chemical shift of adenine N1 shifts downfield, and at sufficiently low pH (<4), the scalar coupling and NOE disappeared. Explain this result. (This experiment was not actually done, but this is the expected result.)
 - With free adenosine, this same chemical shift is observed, but it occurs at much lower pH values. Explain this result.
- 13-7.** The spectrum of pure methanol is shown in (a) below. When HCl is added, the spectrum is that shown in (b) [Z. Luz, D. Gill, and S. Meiboom, *J. Chem. Phys.* **30**, 1540 (1959)].



- a. Explain this result.
- b. The rate of the proton exchange reaction



can be written as $R = k[\text{CH}_3\text{OH}_2^+][\text{CH}_3\text{OH}]$ where the rate constant k has a value of about $10^8 \text{ M}^{-1} \text{ s}^{-1}$. The chemical relaxation time, τ , for this reaction is $\tau = 1/(k[\text{CH}_3\text{OH}_2^+])$. If the above reaction is responsible for the doublet in the spectrum becoming a single peak, estimate the concentration of HCl at which the doublet disappears. Is this consistent with the autoprotolysis constant for methanol: $[\text{CH}_3\text{O}^-][\text{CH}_3\text{OH}_2^+] = 2 \times 10^{-17} \text{ M}^2$?

Applications of Magnetic Resonance to Biology

14.1 INTRODUCTION

Magnetic resonance has been widely applied to biological systems; therefore, the selection of examples to illustrate the information that can be obtained is necessarily quite arbitrary. We first consider some examples of structural determinations utilizing NMR. As previously indicated, the advantage of NMR is that structures in solution can be obtained and variation of these structures in different environments can be readily assessed. The primary drawback of NMR is that only relatively small structures can be easily determined. However, as discussed in Chapter 13 and exemplified in the subsequent sections, new NMR methodology permits the study of very large structures in some cases. In addition, NMR can provide important information about biological systems other than macromolecule structures, as will be illustrated. In the case of ESR, spin labels and paramagnetic ions can provide unique information, especially in unusual environments such as membranes.

14.2 REGULATION OF DNA TRANSCRIPTION

The regulation of transcription is of central importance in biology. We have previously discussed the use of CD to study zinc fingers, a structure of special importance for the interaction of transcription factors with DNA (Chapter 11). In this section, we consider the structure of cyclic AMP response element-binding protein (CBP). It is a large transcriptional adapter protein that mediates transcription responses to intra- and extracellular signals (cf. Refs. 1, 2). This class of proteins has been implicated in the regulation of cell growth, transformation, and differentiation. Defects in this regulator are involved in a multitude of human diseases. The CBP interacts with a variety of transcription factors and other components of transcription regulation to mediate cellular activities.

CBP itself is a very large protein, more than 2000 amino acids, but contains a number of distinct structural and functional domains. Three putative zinc-binding

domains (zinc fingers) were identified by sequence homology. The CH3 domain was selected for study because not only is it a representative zinc finger structure, but it also binds known transcription factors (3). Constructs of varying length were examined, but the final studies were done on an amino acid sequence that was 88 residues long. It contains 13 cysteine and 5 histidine residues. Nine of the cysteine residues were coordinated to Zn^{2+} . This was ascertained from the chemical shifts of $^{13}\text{C}^\beta$ of cysteine. The cysteines that were not coordinated to Zn^{2+} had chemical shifts of 26.5 – 27.1 ppm, whereas the cysteines bound to Zn^{2+} had chemical shifts of 29.2 – 30.7 ppm. The three Zn^{2+} -coordinated histidines also displayed a downfield shift of 8–9 ppm for a ^{13}C ring resonance of the imidazole and an upfield shift for an ^{15}N resonance. Thus, the sites of metal coordination could be determined from one-dimensional NMR. Three Zn^{2+} ions are bound per CH3 domain.

If Zn^{2+} is absent, the NMR spectra indicated a disordered structure. However, in the presence of the metal, a well-defined structure formed. This can be seen in Figure 14-1, where the HSQC spectrum is shown for the protein with three bound Zn^{2+} atoms. Many additional cross peaks can be seen in the presence of Zn^{2+} . The structure was solved using three-dimensional heteronuclear spectra of uniformly

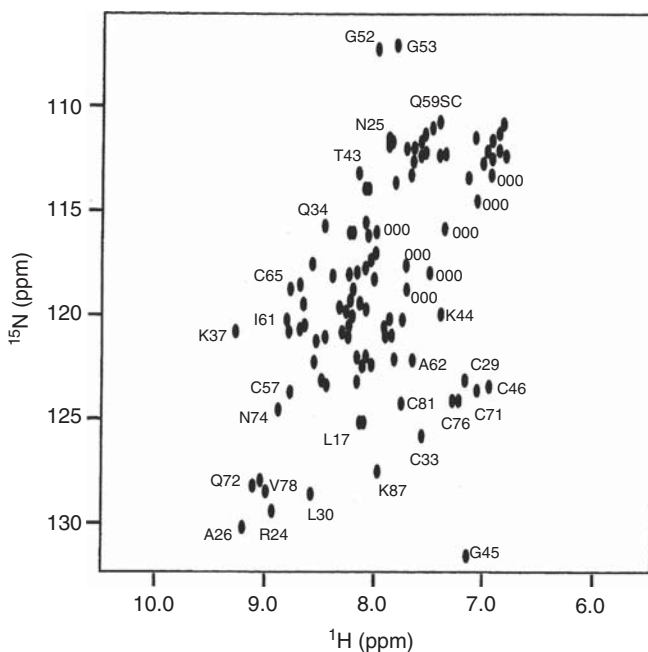


FIGURE 14-1. The 600 MHz proton–nitrogen HSQC spectrum of TAZ2 (CH3) domain, a portion of the c-AMP response element-binding protein, with three bound Zn^{2+} . Many cross peaks can be seen. Some of the amino acid assignments for the resonances are indicated. Reprinted from R. N. De Guzman, H. Y. Liu, M. Martinez-Yamout, H. J. Dyson, and P. E. Wright, Solution Structure of the TAZ2 (CH3) Domain of the Transcriptional Adaptor Protein CBP, *J. Mol. Biol.* **303**, 243 (2000). © 2000, with permission from Elsevier.

labeled ^{15}N and ^{15}N , ^{13}C proteins. The structure determination involved the fitting of NOESY and dihedral angle constraints. In addition, an energy minimization was carried out to arrive at the final structure(s). Energy minimization is a theoretical calculation of the minimum Gibbs energy for the structure, making use of empirical equations for the intramolecular interactions. The final structure is shown in Figure 14-2 in the usual format, namely, the best-fit superposition of multiple structures, in this case 20. In essence, this representation provides a measure of the uncertainty in the final structure. The cysteines and imidazoles are also shown as ball-and-stick representations to indicate how the Zn^{2+} is bound to the protein. Table 14-1 is a brief summary of some of the NMR constraints used to determine the structure and the deviations from ideal covalent geometry.

The structure contains four α -helices and three histidine(cysteine) $_3$ Zn^{2+} binding motifs. The helices are tightly packed to form a hydrophobic core. Two of the metal binding regions are quite similar, but the structure of the third is noticeably different. Although detailed data were not presented, preliminary results suggest that the structure of the CH1 domain is quite similar to that of CH3. Surprisingly, the structures of these zinc fingers are different than others that have been determined, presumably because of unique steric constraints within the structure.

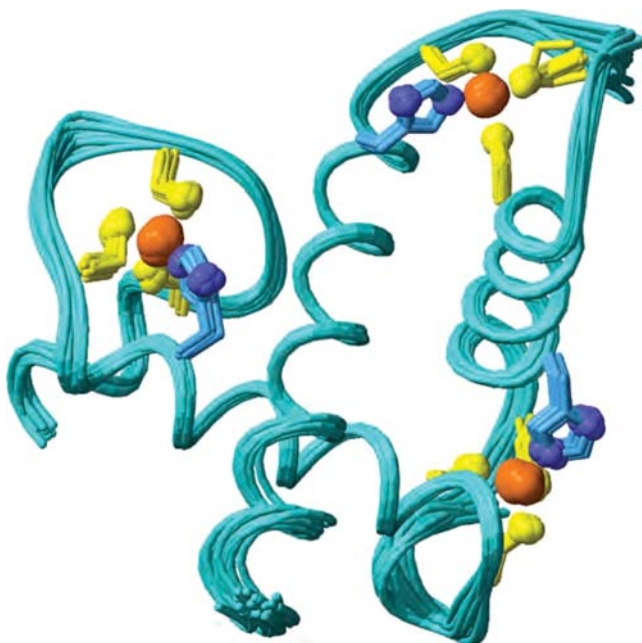


FIGURE 14-2. NMR structure of the TAZ2 (CH3) domain of CBP. A superposition of 20 structures is shown. The orange balls are Zn^{2+} , the yellow ball-and-stick representations are cysteines, and the purple and blue structures are imidazoles from histidines. The four α -helices can be easily discerned. PDB entry 1F81. Copyright by Professor David C. Richardson. Reprinted with permission. Kinemage graphics, then rendered in Raster3D.

TABLE 14-1. NMR Restraints and Statistics for the Structure of the TAZ2 (CH3) Domain

A. NMR restraints	
Total NMR restraints	1030
Distance restraints	846
Intraresidue	145
Sequential	286
Medium range	252
Long range	163
Total dihedral angle restraints	184
B. NOE violations	
Average violation (Å)	0.11 ± 0.04
Maximum violation (Å)	0.25
C. Deviations from ideal covalent geometry	
Bond lengths (Å)	0.0057 ± 0.0001
Bond angles (deg)	2.41 ± 0.02

R. N. De Guzman, H. Y. Liu, M. Martinez-Yamout, H. J. Dyson, and P. E. Wright, *J. Mol. Biol.* **303**, 243 (2000).

As mentioned at the outset, the structure of the CH3 domain is of special interest because it mediates protein–protein interactions that are crucial for regulation of transcription. The interaction with a small peptide (eight residues) from p53, a known activator, with CH3 was investigated with NMR. A specific interaction was found by following shifts in both the backbone and side chains in HSQC spectra. The largest chemical shifts were observed for a small number of residues in three of the helices so that the portion of CH3 interacting with the ligand could be identified. The dissociation constant for the interaction of the peptide and CH3 is about 300 μ M.

This study indicates the variety of information that can be obtained with NMR: the structure of an important biological molecule was determined; the specific residues binding the Zn^{2+} were determined; and the binding region for the interaction of CH3 with a regulatory molecule was identified.

14.3 PROTEIN–DNA INTERACTIONS

Proto-oncogenes are segments of DNA that code for proteins that have a normal function, but can be mutated or altered to become cancer-causing oncogenes. One of these, *c-myc*, encodes a nuclear protein involved in the regulation of transcription. The expression of *c-myc* is in turn regulated by FUSE binding protein (Far-UpStream Element). This protein binds to single-stranded DNA (ssDNA) that is about 1500 base pairs upstream from the *c-myc* promoter. This protein, FBP, contains four homologous repeats (KH) that are separated by linkers of varying lengths (4,5). The minimal ssDNA binding domain is designated as KH3–KH4. The structure of the KH3–KH4

domain bound to ssDNA (20 – 29 nucleotides long) has been solved using multidimensional NMR (6). The total molecular weight of the complex is about 30,000, and 3153 NMR restraints were used to determine the structure. A representation of the best-fit structure for the KH3–ssDNA is shown in Figure 14-3. The ssDNA is shown as a ball-and-stick model, whereas the protein chain is shown to delineate the overall fold and the α -helices and β -sheets.

The protein fold has three α -helices packed onto a three-stranded antiparallel β -sheet. The ssDNA binds in a groove formed by helices 1 and 2. The center of the groove is hydrophobic and the edges are hydrophilic so that the DNA bases point toward the center and the sugar phosphates toward the left-hand side of the protein. As expected, recognition of the DNA involves a number of intermolecular hydrogen bonds. This protein binding scaffold for nucleic acid binding can be fine-tuned for either ssDNA or RNA. The KH3 and KH4 domains do not interact with each other. The flexible linker between the domains is 30 amino acids long. NMR relaxation time measurements indicate that the two domains wobble with respect to each other, with time constants of nanoseconds. The flexibility of the domains appears to be of functional significance. If the flexibility of the motion is restricted by deleting four of the DNA bases in the intervening ssDNA between domains, the *c-myc* expression is reduced by a factor of four, whereas adding four bases has no effect.

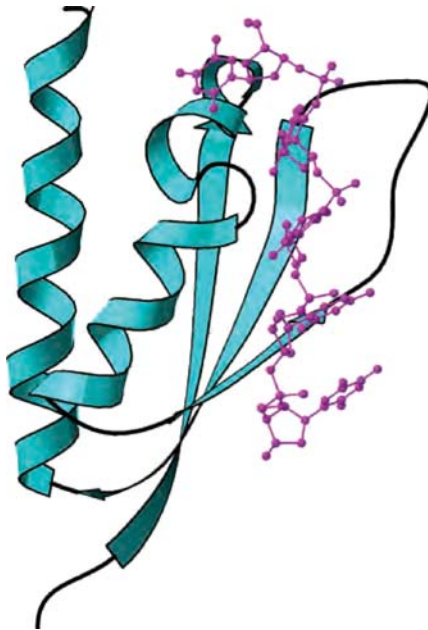


FIGURE 14-3. Structure of a KH3–DNA complex determined with NMR. The DNA (10 mer) is shown in magenta as a ball-and-stick model, and the protein backbone is in cyan. PDB entry 1J5K. Copyright by Professor David C. Richardson. Reprinted with permission. Kinemage graphics, then rendered in Raster3D.

What is the significance of this structure? The functional concept is that DNA transcription can be controlled at a significant distance along the DNA by recognition of the ssDNA that is formed during transcription. The structure determined shows that the regulatory element (FUSE) upstream from the *c-myc* promoter is recognized specifically and forms a very stable complex with the FBA. This work, when combined with functional studies, can be used to construct a mechanism for transcriptional control. Furthermore, the overexpression of *c-myc* has been linked to cancer: the loss of FBP prevents *c-myc* expression and halts cellular proliferation. The disruption of the FBP–ssDNA interaction, therefore, is a potential target for cancer therapy.

14.4 DYNAMICS OF PROTEIN FOLDING

Understanding how proteins fold into their native conformation is of central importance in biology. First, protein folding is a necessary part of cellular metabolism and development. Second, understanding how proteins fold provides information about the intramolecular forces in proteins. Many proteins fold and unfold very fast, in times less than a second, so that study of the dynamics of protein folding requires special methods, including the use of NMR (cf. Ref. 7). Protein folding has been previously discussed in Chapters 3 and 11.

A prototypical fast folding protein is the N-terminal domain of bacteriophage λ repressor, a crucial molecule in gene regulation (cf. Ref. 8). The structure of the N-terminal domain has been determined with both crystallography and NMR, and it is essentially identical to the full-length version of λ repressor. The protein unfolds in a simple two-state process, both thermally and in the presence of urea. The folding reaction can be written as



where N and D are the native and denatured states, respectively, and k_f and k_u are the first-order rate constants for folding and unfolding, respectively.

The aromatic region of the NMR spectrum is quite different for the native and denatured states as shown in Figure 14-4, where the NMR spectra are shown at various urea concentrations (9–11). Both the native and denatured states show well-resolved spectra. Some of the central peaks are in fast exchange at all urea concentrations so that only sharp resonances are observed, and the resonance most downfield is in slow exchange. However, some of the peaks show broadening at intermediate urea concentrations because the rate of interconversion of the native and denatured states is comparable to the chemical shift between the native and denatured resonances, about 100 cps.

The resonances chosen for analysis are associated with two specific tyrosine residues. The analysis of the data calculated the line shapes from theoretical considerations and compared the calculated and experimental line shapes with varying rates of reaction until the two matched. Basically, this measured the effect of the reaction rate on the spin–spin relaxation time, T_2 , and yielded the relaxation time for the reaction in Eq. (14-1), τ (Eq. (13-8)). The reciprocal of the relaxation time for the

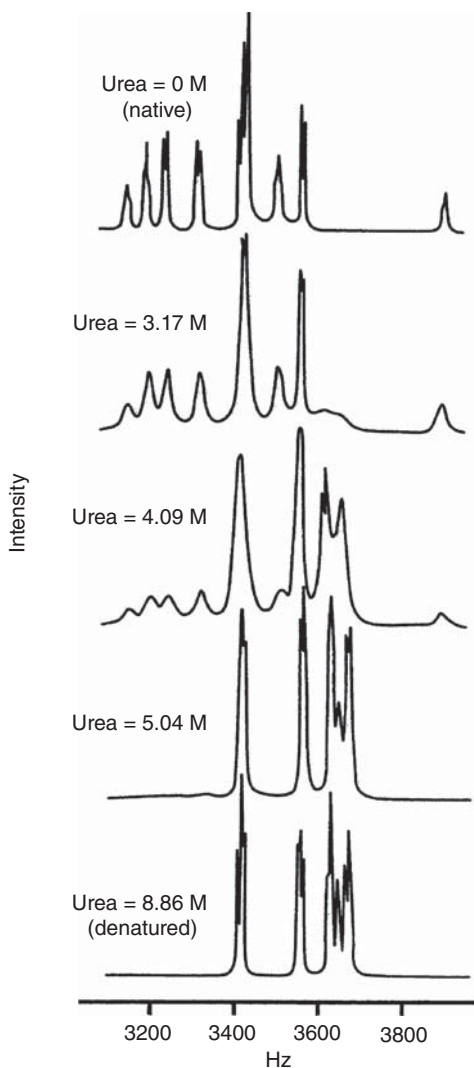


FIGURE 14-4. NMR spectra of the aromatic region of λ repressor at various concentrations of urea. Differences between the native and denatured states (top and bottom) can be seen. The sharp peaks in the middle are in fast exchange. At intermediate urea concentrations, line broadening of several of the peaks can be seen because of chemical exchange. Analysis of the line broadening permitted determination of the rate constants for the interconversion of the native and denatured proteins. Reprinted in part with permission from J. K. Myers and T. G. Oas, *Biochemistry* **38**, 6761 (1999). ©1999 by American Chemical Society.

chemical reaction is the sum of the two rate constants (Chapter 5). The proportion of the protein present in the native and denatured states can be derived from the line shape analysis as it is directly proportional to the strength of the resonance for each state. Thus, the individual rate constants can be directly determined. The rate constants vary linearly with the urea concentration but are in the range of $10^2 - 10^3 \text{ s}^{-1}$. When extrapolated to zero urea concentration, the rate constants for folding and unfolding are 3600 s^{-1} and 27 s^{-1} at 37°C , respectively, for the particular variant of the λ repressor studied (9).

Why does this protein fold so fast? It is relatively small, only 79 amino acid residues, but some proteins of this size fold much slower. A unique feature of this protein is that the only significant secondary structural elements are α -helices, which are known to wind and unwind in microseconds or less. In contrast, β -sheet structures form and break down considerably slower. The detailed folding mechanism has been explored further by site-specific mutagenesis of amino acid residues. The results obtained suggest that minor modifications, such as changing two glycines to alanines or disruption of a single hydrogen bond, can have significant effects on the folding mechanism. The results further suggest that formation of one of the five helical stretches may be a crucial slow step in the folding mechanism.

Knowledge of the kinetics of fast folding proteins provides unique information about the folding mechanism, and NMR provides one of the few methods available for studying these very fast reactions.

14.5 RNA FOLDING

The folding of RNA into functional structures is of obvious importance to biology. The mechanism of folding appears to be more complex than that for proteins (cf. Ref. 12). A number of structures exist that have similar Gibbs energies so that molecules can become trapped in nonfunctional structures. Furthermore, because RNA molecules are highly charged, the structures formed are very dependent on the ionic environment, particularly on the concentration of Mg^{2+} . A frequently invoked mechanism is that two major structural changes occur in the folding reaction. First, stable secondary structures form on the microsecond time scale. Secondary structure is defined as local structural elements such as helices, due to base pairing. The second stage is formation of tertiary folding that brings the secondary structural elements together. This is similar to the mechanism of protein folding discussed earlier. The study described subsequently suggests that this picture is too simple in many cases.

We have previously discussed ribozymes (Chapters 6 and 9). The folding of a self-splicing RNA from *Tetrahymena* has served as a prototypical example of RNA folding (13,14). This group I intron consists of two large domains, labeled P4–P6 and P1–P2/P3–P9. The stable P4–P6 domain folds independently, and its crystal structure has been determined by X-ray crystallography (15,16). Within this structure are three helices, labeled P5abc, and the structure of this RNA has been studied with NMR (17). The P5abc RNA can fold independently, and its structure is amenable to determination by NMR.

To arrive at a structure, the one-dimensional proton NMR spectrum was obtained for the imino protons, which are involved in the hydrogen bonding of base pairs. An A–U base pair has one imino proton, and a G–C base pair has two imino protons. A two-dimensional NOESY study was then carried out to determine which imino protons were close to each other. Interpretation of these spectra required assignment of each of the resonance peaks in the one-dimensional spectrum to specific bases in the RNA. The imino protons of guanine and uracil give rise to very sharp NMR resonances in the 10–15 ppm range. This suggests that a single conformation is present and no aggregation is occurring. In the NOESY spectrum, the diagonal peaks corresponded to the one-dimensional spectrum, and each cross peak connected two diagonal peaks, which implies that the two protons are within 5 Å of each other. Neighboring base pairs give rise to NOEs and helped establish the assignments of specific resonance peaks and the secondary structure of the RNA.

The results are summarized in Figure 14-5, where the secondary structure of the 56-nucleotide RNA is shown. The arrows indicate the connectivity established by the NOESY experiments, often called an “NOE walk.” Included in the structure are disks that indicate the locations of the imino protons, all of which are involved in base pairs.

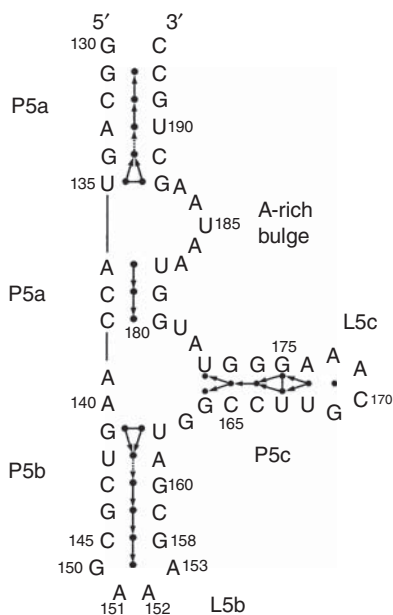


FIGURE 14-5. Secondary structure of the 56-nucleotide RNA ribozyme fragment. The disk between each base pair represents the imino protons that are observed by NMR. The arrows represent the connectivity established by the NOESY experiments. The dotted arrow NOE was not observed because the resonances could not be resolved. Reproduced with permission from M. Wu and I. Tinoco, Jr., *Proc. Natl. Acad. Sci. USA* **95**, 11555 (1998). © 1998 National Academy of Sciences, USA.

The NMR results give a quite clear picture of the secondary structure. Surprisingly, when this structure was compared with that determined with X-ray crystallography, the two structures were found to be different. In the crystals, the hydrogen bonding of six base pairs is broken (two G–C, one A–U, and three G–U) and four new G–C base pairs are formed, as well as two nonstandard A–U pairs. In addition, a tetraloop is disrupted, and other changes occur. The calculated Gibbs energy of the secondary structure of the RNA crystal structure is actually higher than that of the NMR secondary structure.

However, this difference in Gibbs energy is more than compensated for by interactions within the tertiary structure. These structural differences can be reconciled if Mg^{2+} is added to the solution. The addition of this metal results in a conversion to a new structure that is identical to that found in the crystal structure. In fact, two sets of resonances are seen, indicating that the two structures are slowly interconverting. In this case, slow means with a rate constant less than 30 s^{-1} . (This is determined by the differences in chemical shifts between the two structures.) In terms of the structure, the interactions of RNA with Mg^{2+} alter both the secondary structure and the tertiary structure. This change in structure is shown schematically in Figure 14-6. In the

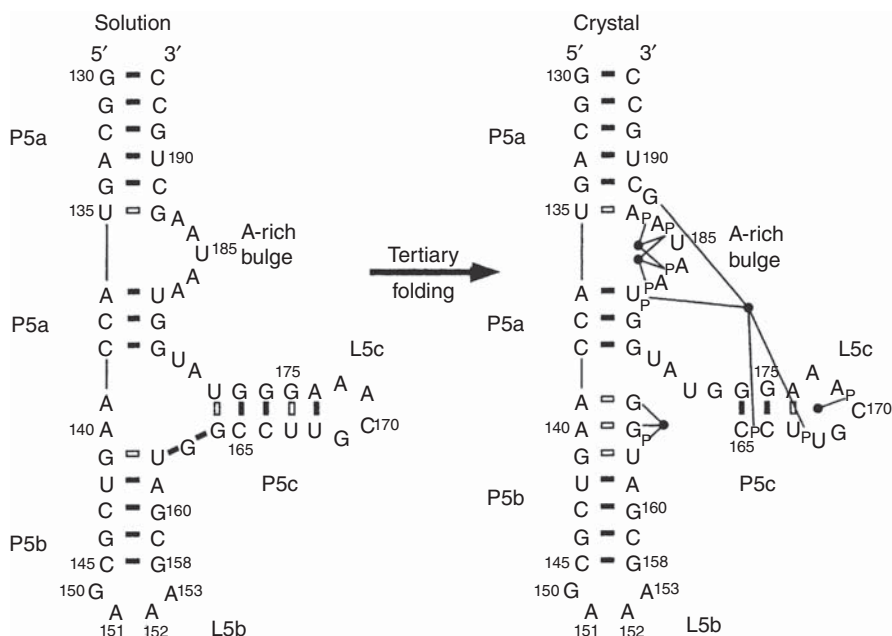


FIGURE 14-6. Comparison of the structures of the RNA ribozyme fragment determined by NMR and crystallography. The tertiary folding is caused by the five Mg^{2+} ions shown in the crystal structure as filled circles. The solid bars between bases represent Watson–Crick pairing, and the open bars represent non-Watson–Crick base pairing. Note the changes in base pairing caused by the presence of Mg^{2+} . Reproduced with permission from M. Wu and I. Tinoco, Jr., *Proc. Natl. Acad. Sci. USA* **95**, 11555 (1998). © 1998 National Academy of Sciences, USA.

fully folded RNA, a cluster of five Mg^{2+} ions are coordinated to phosphate oxygens, indicating that the tertiary folding occurs on a Mg^{2+} core.

The significance of these results in understanding RNA folding is that the pre-formed secondary structure cannot be assumed to be the direct precursor of the final tertiary structure (17,18). In other cases, this has been found to be a valid assumption. It also indicates that many stable conformations are available for RNA within a range of salt and Mg^{2+} concentrations. This makes the delineation of the physiological folding mechanism for RNA more difficult than that for proteins, where a more restricted number of stable conformations probably exists.

In a follow-up study, site-specific mutations of the P5abc RNA fragment were made to explore the balance between tertiary and secondary structures (19). In particular, point mutations designed to disrupt the secondary structure greatly affected the Mg^{2+} -dependent folding into the new structure. Moreover, two single-point mutations are sufficient to prevent the rearrangement of the secondary structure observed in the crystal structure. However, if the same experiments are carried out with the P4–P6 domain, formation of the tertiary structure is sufficient to alter the secondary structure of P5abc, as observed with the native P5abc RNA, even with the point mutations. The message of this work is that the balance between tertiary and secondary structure is very delicate in RNA. It depends on both the ionic conditions and the context of the RNA domains. This is likely to be a general feature of RNA structure/folding.

14.6 LACTOSE PERMEASE

Transport proteins are integral membrane proteins that are responsible for the flow of many metabolites across the cell membrane. Two broad classes of transport mechanisms exist, *facilitated diffusion* and *active transport*. Facilitated diffusion depends on a concentration gradient, and the molecules being transported flow from a higher to a lower concentration. Selectivity is due to the size of the pore and/or gating of the channel by a stimulus. For example, Gram-negative bacteria contain several porins, 34–38 kDa proteins, in their outer membranes. The porins permit molecules and ions with a molecular mass of about 600 to enter. The molecules that are transported include maltodextrins, sugar phosphates, and chelated iron. Glucose and HCO_3^- are also generally transported across membranes by facilitated diffusion. In active transport, material is carried across the membrane against a concentration gradient, from low to high concentration. Active transport must be coupled to an electrochemical gradient created by a free energy favorable process such as the hydrolysis of ATP.

Lactose permease (LacY) is responsible for all of the translocation reactions carried out by the galactoside transport system in *E. coli*. It couples the Gibbs energy associated with the collapse of a proton gradient, that is, transport of protons, with the energetically uphill translocation of galactosides against a concentration gradient. The lactose permease from *E. coli* has been extensively studied (cf. Ref. 20), and ESR has been one of the tools used to elucidate its structure in the membrane. The protein has been purified, reconstituted into proteoliposomes, and shown to be solely responsible for the galactoside transport.

The crystal structure of lactose permease is now known (21), but prior to the structure determination, the primary features of the molecule had been determined by less direct methods such as mutagenesis, chemical cross-linking, and ESR. Lactose permease has a molecular weight of about 45,500, and the structure contains 12 α -helices passing through the membrane, linked by loops on each side of the membrane. The structure is shown schematically in Figure 14-7. A hydrophilic cavity is formed between the helices that alternately faces the inside and outside of the cell as the sugar is transported. We will not be concerned about the detailed mechanism of transport here, but rather with the structural information that was obtained from a series of ESR studies. In simplistic terms, the mechanism consists of a series of sugar and proton bindings that result in the conformation of the protein switching between conformations that expose the hydrophilic cavity to the appropriate side of the membrane. The structural models for the inward- and outward-facing conformations are shown schematically in Figure 14-8: note the complex arrangement and distortions of the helical rods that make up the structure.

Site-directed spin labeling of lactose permease can be carried out by introduction of a single cysteine residue into a protein with no cysteines, and then labeling the thiol with a nitroxide spin label. The protein with no cysteines (prepared by site-specific mutagenesis) retains its transport activity. In the initial set of experiments (22), three derivatives were prepared with cysteines at amino acid positions 148 (helix V) and 228 (helix VII), 148 and 226 (helix VII), or 148 and 275 (helix VIII). A nitroxide spin label was then covalently linked to the cysteines. The ESR spectra showed relatively broad lines for all three pairs, suggesting that the nitroxides were relatively immobile due to strong interactions with their environment, nearby protein and membrane. Additional broadening could be detected at low temperatures for the 148/228 and 148/275 pairs due to close proximity of the nitroxides (spin-spin interactions). Analysis of this additional broadening indicates that the spin labels are within about 15 Å

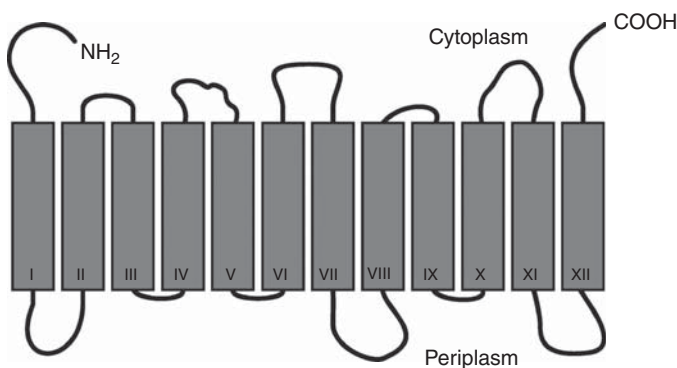


FIGURE 14-7. Secondary structure model of lac permease. The permease has a hydrophilic N-terminus, followed by 12 α -helical hydrophobic domains that are connected by hydrophilic loops, and a hydrophilic C-terminus. Adapted with permission from J. Wu, J. Voss, W. L. Hubbell, and H. R. Kaback, *Proc. Natl. Acad. Sci. USA* **93**, 10123 (1996). © 1996 National Academy of Sciences, USA.

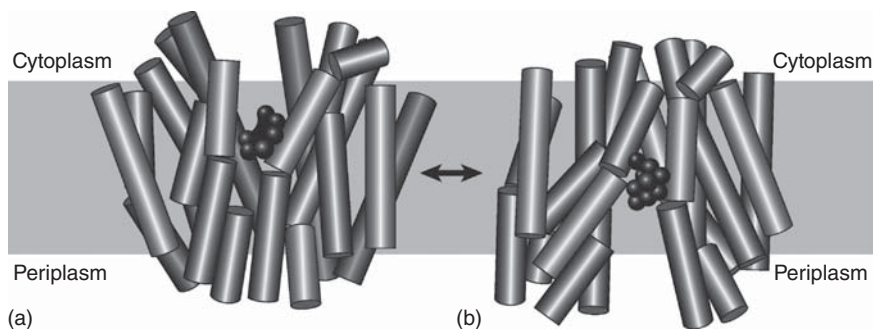


FIGURE 14-8. Schematic representation of the inward and outward-facing conformations of lac permease. The mechanism of lactose transport involves alternating these conformations. Lactose is shown as a space-filling molecular structure. Adapted with permission from J. Abramson, I. Smirnova, V. Kasho, G. Verner, H. R. Kaback, and S. Iwata, *Science* **301**, 610 (2003). © 2003 AAAS.

of each other. These results suggest that position 148 in helix V is in close proximity to position 228 in helix VII and position 275 in helix VIII. Chemical cross-linking experiments indicate that position 148 is closer to helix VII than to helix VIII. On the other hand, the lack of spin–spin interactions between positions 148 and 226 suggests that cysteine 226 is on the opposite face of helix VII from cysteine 228.

In a second study (23), a high-affinity Cu^{2+} binding site was created on lactose permease by replacing Arg 301 (helix IX) and Glu 325 (helix X) with His residues. In addition, a series of proteins with single cysteines at various positions in helices II, V, or VII were prepared, and nitroxide spin labels were attached to the cysteines. The ESR spectra of the nitroxides varied, depending on their locations. Those interacting within the protein structure were broad, indicating restricted rotation, whereas those directed away from the interior of the protein were much sharper, indicating freer rotation. When Cu^{2+} was added, some of the lines broadened due to interaction of the unpaired electron of the metal ion with the spin label. This permitted a mapping of the cysteines with respect to the metal ion. When these results were combined with those from chemical cross-linking, fluorescence, and other techniques, a model for the helix packing could be developed. Additional experiments of this type utilized the interaction of Gd(III) with spin labels to provide confirmation that helix V lies close to both helices VII and VIII (24).

Single cysteines also were introduced at positions 126–156, and the cysteines were derivatized with nitroxide spin labels (25). The dynamics of these labels varied from highly mobile to highly immobilized, as judged by the broadness of the spectra. These results can be interpreted in terms of structure in that enhanced mobility can be interpreted as increased accessibility of the side chain. In addition, spectral broadening due to addition of paramagnetic species (potassium chromium oxalate and oxygen) provided further information about the accessibility of the spin labels to the solvent. These experiments provide information about both the structure and the dynamics of the protein side chains.

These extensive ESR studies provide detailed structural information about a membrane-bound protein. Such information is difficult to obtain because membrane-bound proteins are not readily accessible to the large cadre of techniques that can be applied to soluble proteins. In this case, a crystal structure has been obtained that is in reasonable accord with the more indirect methods used. The studies of lactose permease serve as a useful paradigm for the investigation of the structure of membrane-bound proteins.

14.7 PROTEASOME STRUCTURE AND FUNCTION

As a final example of an important biological system studied with NMR, the *proteasome* is considered. The proteasome is a molecular device that removes damaged and misfolded proteins from eukaryotes and archaea. It is critical for the regulation of important cellular processes, including cell division and signal transduction. The proteasome is a megadalton structure consisting of many different polypeptide chains. NMR studies have been carried out on a 670,000 dalton 20S core particle from *Thermoplasma acidophilum* (26–28). It consists of four homoheptameric rings made up of two polypeptide chains, α and β , that are stacked to form three interconnected cavities. Two of the cavities are antechambers, $\alpha_7\beta_7$, which the substrates pass through to reach the chamber in which peptide bond hydrolysis occurs, $\beta_7\beta_7$. Access to the catalytic chamber occurs through gated pores at each end of the catalytic chamber.

The 20S proteasome is too large to study by conventional NMR, but this limitation was overcome by labeling isoleucine, leucine, and valine methyl groups with protons in an otherwise deuterated background. Sprangers and Kay (26) were able to assign NMR resonances to approximately 90 of the methyl groups in the α polypeptide chains using TROSY methodology. Some of the residues inside the antechamber providing access to the catalytic chamber were found to undergo correlated motions on the millisecond time scale that may be related to passage of the substrate through the proteasome. A molecular gating system was suggested on the basis of these experiments.

A subsequent study (27) examined protein substrates within the antechambers, also using TROSY. The actual proteolysis of protein substrates occurs on unstructured proteins rather than on proteins in their native structures. Three different substrate proteins were shown to interact with the antechamber walls so that they were maintained in unstructured states. Thus the antechambers alter the properties of the substrates so that they exist in an unstructured state in order to provide a substrate that can be readily hydrolyzed.

As might be expected, the activity of the proteasome is highly regulated. The core proteasome exists in multiple conformations that are readily interconverted. The 11S activator binds about 75 Å from the catalytic site so that the activation of proteolysis must be indirect. NMR studies of the binding of the 11S activator to the core proteasome demonstrated that changes in the relative populations of the proteasome conformers occur upon binding and alter proteolysis patterns (28). This suggests that changes in protein conformations are responsible for regulation of the activity of the

proteasome. Regulation of this type (allostery) is discussed in Chapter 17. This elegant work demonstrates that NMR can provide unique information about very large molecular complexes by overcoming resolution limitations with special methods.

14.8 CONCLUSION

These selected examples of the application of magnetic resonance methods to biology are a small fraction of the many interesting investigations that have been carried out. Note that in each case, important questions were being asked about structure and function. In point of fact, the biological problem is the most significant aspect of these studies, and magnetic resonance is one of many spectroscopic/biochemical tools that can be used for the elucidation of these problems.

REFERENCES

1. A. Giordano and M. L. Avantaggiati, *J. Cell Phys.* 181, 218 (1999).
2. R. H. Goodman and S. M. Smolik, *Genes Dev.* 14, 1553 (2000).
3. R. N. De Guzman, H. Y. Liu, M. Martinez-Yamout, H. J. Dyson, and P. E. Wright, *J. Mol. Biol.* 303, 243 (2000).
4. R. Duncan, L. Bazar, G. Michelotti, T. Tomonaga, H. Krutzsch, M. Avigan, and D. Levens, *Genes Dev.* 8, 465 (1994).
5. C. G. Burd and G. Dreyfuss, *Science* 265, 615 (1994).
6. D. T. Braddock, J. M. Louis, J. L. Baber, D. Levens, and G. M. Clore, *Nature* 415, 1051 (2002).
7. J. K. Myers and T. G. Oas, *Annu. Rev. Biochem.* 71, 783 (2002).
8. R. T. Sauer, S. R. Jordan, and C. O. Pabo, *J. Mol. Biol.* 227, 177 (1990).
9. G. S. Huang and T. G. Oas, *Proc. Natl. Acad. Sci. USA* 92, 6878 (1995).
10. R. E. Burton, G. S. Huang, M. A. Daugherty, P. W. Fullbright, and T. G. Oas, *J. Mol. Biol.* 263, 311 (1996).
11. J. K. Myers and T. G. Oas, *Biochemistry* 38, 6761 (1999).
12. P. Brion and E. Westhof, *Annu. Rev. Biophys. Biomol. Struct.* 26, 113 (1997).
13. P. P. Zarrinker and J. R. Williamson, *Nat. Struct. Biol.* 3, 432 (1996).
14. B. Scalvi, M. Sullivan, M. R. Chance, M. Brenowitz, and S. A. Woodson, *Science* 279, 1940 (1998).
15. J. H. Cate, A. R. Gooding, E. Podell, K. Zhou, B. L. Golden, C. E. Kundrot, T. R. Cech, and J. A. Doudna, *Science* 273, 1678 (1996).
16. J. H. Cate, A. R. Gooding, E. Podell, K. Zhou, B. L. Golden, A. A. Szewczak, T. R. Cech, and J. A. Doudna, *Science* 273, 1696 (1996).
17. M. Wu and I. Tinoco, Jr., *Proc. Natl. Acad. Sci. USA* 95, 11555 (1998).
18. D. Thirumalai, *Proc. Natl. Acad. Sci. USA* 95, 11506 (1998).
19. S. K. Silverman, M. Zheng, M. Wu, I. Tinoco, Jr., and T. R. Cech, *RNA* 5, 1665 (1999).
20. J. Abramson, I. Smirnova, V. Kasho, G. Verner, S. Iwata, and H. R. Kaback, *FEBS Lett.* 555, 96 (2003).

21. J. Abramson, I. Smirnova, V. Kasho, G. Verner, H. R. Kaback, and S. Iwata, *Science* 301, 610 (2003).
22. J. Wu, J. Voss, W. L. Hubbell, and H. R. Kaback, *Proc. Natl. Acad. Sci. USA* 93, 10123 (1996).
23. J. Voss, W. L. Hubbell, and H. R. Kaback, *Biochemistry* 37, 211 (1998).
24. J. Voss, J. Wu, W. L. Hubbell, V. Jacques, C. F. Meares, and H. R. Kaback, *Biochemistry* 40, 3184 (2001).
25. M. Zhao, K. C. Zen, J. Hernandez-Borrell, C. Altenbach, W. L. Hubbell, and H. R. Kaback, *Biochemistry* 38, 15970 (1999).
26. R. Sprangers and L. E. Kay, *Nature* 445, 618 (2007).
27. A. M. Ruschak, T. L. Religa, S. Breuer, S. Witt and L. E. Kay, *Nature* 467, 868 (2010).
28. A. M. Ruschak and L. E. Kay, *Proc. Natl. Acad. Sci. USA* 109 E3454 (2012).

STATISTICAL MECHANICS

Fundamentals of Statistical Mechanics

15.1 INTRODUCTION

Statistical mechanics provides a bridge linking microscopic molecular properties and macroscopic thermodynamic properties. Equilibrium statistical mechanics deals with the average behavior of large numbers of molecules. In this chapter, we start by introducing the kinetic model of gases, which relates the microscopic properties of an ideal gas to macroscopic properties such as temperature and pressure. An important outcome of this discussion is an expression relating the average molecular speed to the temperature of the system. We then move on to the topic of statistical mechanics, initially considering systems comprised of noninteracting molecules but later extending this treatment to describe more realistic systems with interacting molecules. This section includes a derivation of the Boltzmann distribution, which relates the populations of molecular states to the energies of these states. Along the way, we also define new concepts that are central to statistical mechanics, such as partition functions and ensembles. The partition function is a measure of the thermally accessible states in a system and contains thermodynamic information required to calculate quantities such as internal energy and entropy. This discussion also provides a statistical definition of entropy, which is consistent with our previous thermodynamic definition of entropy but provides additional insights at the molecular level.

The concepts of statistical mechanics are directly applicable to biological systems. The last section of this chapter applies these ideas to the helix-coil transition in polypeptides. In Chapter 16, the fundamental concepts of statistical mechanics provide the framework for understanding computational methods developed for studying biological systems. Many of the ideas discussed in this chapter, such as the relation between temperature and velocities, the molecular description of entropy, and the definitions of statistical ensembles, are central to the simulation of protein folding, protein dynamics, and a wide range of other biological processes.

15.2 KINETIC MODEL OF GASES

The kinetic model of gases allows us to understand macroscopic properties such as pressure and temperature on a microscopic level. Many of the concepts developed

in this section will be used in the subsequent sections on statistical mechanics. The kinetic model is based on three assumptions. First, the gas consists of molecules of mass m that are moving continuously and randomly. Second, the size of the molecules is negligible compared to the average distance traveled between collisions. Third, the molecules interact only through brief, infrequent, and elastic collisions, which conserve the total translational kinetic energy of the molecules.

We start by considering an ideal gas in a large container and calculating the pressure exerted by the gas on the walls of the container. We assume that this pressure, P , is due to collisions of the molecules with the walls and use Newton's equations of motion to determine the pressure, which is defined to be the force per unit area. The resulting pressure for n moles of the ideal gas in a box of volume V is:

$$PV = \frac{1}{3}nM\langle v^2 \rangle \quad (15-1)$$

Here $M = N_A m$ is the molar mass of the molecules, where N_A is Avogadro's number. Moreover, $\langle v^2 \rangle$ is the mean-square speed, defined as the average of the squares of the speeds of the molecules, where the speed v is simply the magnitude of the velocity vector. Eq. (15-1) relates the pressure that the molecules exert on the walls of the container to the mean-square speed of the molecules.

According to the ideal gas law, the product of the pressure and volume is proportional to the temperature T of the system:

$$PV = nRT \quad (15-2)$$

where R is the gas constant. Combining Eqs. (15-1) and (15-2), we obtain the following expression for the mean-square speed:

$$\langle v^2 \rangle = \frac{3RT}{M} \quad (15-3)$$

Now we have related the mean-square speed of the molecules to the temperature of the system. This expression is often expressed in terms of the root-mean-square speed c , given as

$$c = \langle v^2 \rangle^{1/2} = \left(\frac{3RT}{M} \right)^{1/2} \quad (15-4)$$

Thus, the root-mean-square speed of the molecules is proportional to the square root of the temperature and inversely proportional to the square root of the molar mass of the molecules. From a physical perspective, this equation means that molecules tend to move faster at higher temperatures, and heavier molecules tend to move more slowly than lighter molecules at a given temperature. However, this equation only tells us about average quantities and provides no information about the speeds of individual molecules.

In general, the speeds of the individual molecules vary over a wide range, and collisions continually redistribute the speeds among the different molecules. For example, a specific molecule may slow down or speed up after a collision. Because we are

interested in macroscopic properties, we focus on characterizing the distribution of the molecular speeds rather than determining the speeds of individual molecules. For this purpose, we define a distribution function $f(v)$, where $f(v)dv$ represents the fraction of molecules with speed between v and $v + dv$. The form of this distribution function was derived by Maxwell to be

$$f(v) = 4\pi \left(\frac{M}{2\pi RT} \right)^{3/2} v^2 e^{-Mv^2/2RT} \quad (15-5)$$

This distribution function is called the Maxwell distribution of speeds and is depicted in Figure 15-1.

The Maxwell distribution of speeds is the product of the square of the speed, v^2 , and an exponential function that depends on the square of the speed. The prefactor, v^2 , causes the distribution function to approach zero as the speed approaches zero, and the exponential function causes the distribution function to approach zero as the speed approaches infinity. The dependence of the exponential factor on both the molar mass M and the temperature T causes this exponential to decay to zero more rapidly for higher molar mass and lower temperature. Thus, higher speeds are more probable for lighter molecules and higher temperatures. These trends are illustrated by the distribution functions shown in Figure 15-1.

To determine the fraction of molecules with speeds between v_1 and v_2 , we calculate the following integral:

$$\text{Fraction in range } v_1 \text{ to } v_2 = \int_{v_1}^{v_2} f(v)dv \quad (15-6)$$

The fraction of molecules with speeds in this range can also be interpreted as the probability that a molecule will have a speed in this range. This probability is illustrated by the shaded region in Figure 15-2, which represents the area under the distribution

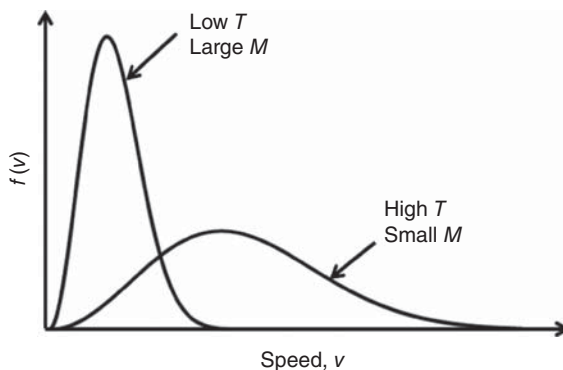


FIGURE 15-1. Maxwell distribution of speeds, as given by $f(v)$ in Eq. (15-5). The distribution function is given for two different values of temperature T and molar mass M . These curves show that higher speeds are more probable for lighter molecules and higher temperatures.

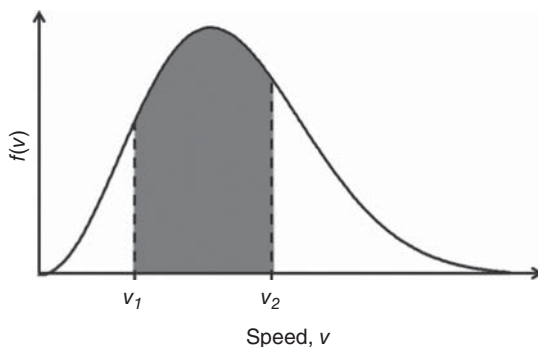


FIGURE 15-2. Maxwell distribution of speeds, denoted by $f(v)$. The fraction of molecules with speeds between v_1 and v_2 , as given by the integral in Eq. (15-6), is indicated by the shaded region. This shaded region can also be viewed as the probability that a molecule will have a speed in this range.

function $f(v)$ between v_1 and v_2 . The integral of $f(v)$ over all speeds, ranging from zero to infinity, is unity because all molecules would be included in this integral. In other words, the area under the entire distribution function curve is unity.

The average speed, or mean speed, can be calculated by summing over the speeds of all molecules and dividing by the total number of molecules. However, we do not know the speeds of the individual molecules and instead need to use the distribution function for this purpose. To understand this procedure, first we consider a system in which the molecules have discrete speeds, v_i , and we determine the fraction of molecules, p_i , that have speed v_i . The fraction of molecules p_i can also be viewed as the probability that a given molecule will have speed v_i . For this discrete system, the average speed can be calculated by summing over all possible speeds, multiplying each value by the probability of that speed occurring. Mathematically, this summation is expressed as

$$\langle v \rangle = \sum_i p_i v_i \quad (15-7)$$

This summation is equivalent to summing over the speeds of all molecules and dividing by the total number of molecules.

To understand this equivalence, consider a group of six students who took a quiz worth three points. Three of the students obtained a score of 1, two of the students obtained a score of 2, and one student obtained a perfect score of 3. (Perhaps some of these students should have studied a little harder!) To calculate the average score, we can add up all of the scores and divide by the number of students: $(1+1+1+2+2+3)/6 = 5/3$. Alternatively, we can calculate the fraction of students who obtained each score: $1/2$ of the students obtained a score of 1, $1/3$ of the students obtained a score of 2, and $1/6$ of the students obtained a score of 3. Now we can calculate the average score by summing over all possible scores, multiplying each score by the fraction of students who received that score: $[(1/2) \times 1 + (1/3) \times 2 + (1/6) \times 3] = 5/3$. It is clear that these two averaging procedures lead to the same

result. This example clarifies that the expression in Eq. (15-7) is simply another way of calculating the average of the speeds of the molecules.

Now assume that the speeds can have continuous values, as in the real system. In this case, the summation becomes an integral, and the probability is given by $f(v)dv$ instead of p_i . In this case, the average speed can be calculated as

$$\langle v \rangle = \int_0^{\infty} v f(v) dv = \left(\frac{8RT}{\pi M} \right)^{1/2} \quad (15-8)$$

Similar to the root-mean-square speed given in Eq. (15-4), the average speed is proportional to the square root of the temperature and inversely proportional to the square root of the molar mass of the molecules. However, the constants of proportionality are different for these two different types of speeds.

We can also calculate the most probable speed, c^* , which is the speed at which the distribution function has a maximum. This speed can be determined by differentiating $f(v)$ with respect to v and finding the value of v for which this derivative is zero. Setting the derivative of $f(v)$ to zero leads to

$$\frac{df(v)}{dv} = 4\pi \left(\frac{M}{2\pi RT} \right)^{3/2} \left[2v - \frac{Mv^3}{RT} \right] e^{-Mv^2/2RT} = 0 \quad (15-9)$$

and solving for the speed v in this equation leads to the most probable speed:

$$c^* = \left(\frac{2RT}{M} \right)^{1/2} \quad (15-10)$$

Thus, we have derived equations for the root-mean-square speed (Eq. (15-4)), the average speed (Eq. (15-8)), and the most probable speed (Eq. (15-10)). In general, $c^* < \langle v \rangle < c$, so the root-mean-square speed and the average speed are greater than the speed associated with the maximum of $f(v)$. Moreover, all of these types of speed increase as the temperature increases and as the molar mass decreases.

In addition to the distribution of molecular speeds, we can also calculate the distribution of molecular translational energies, denoted $g(\epsilon)$. These two distribution functions are related because the molecular translational energy, which is simply the kinetic energy, is $\epsilon = mv^2/2$. Thus, we can convert $f(v)$ to $g(\epsilon)$ through a simple variable transformation from speed to energy. Using this expression for the kinetic energy, $v^2 = 2\epsilon/m$ and $d\epsilon = mv dv = (2m\epsilon)^{1/2} dv$. This latter relation can be rearranged to obtain $dv = (2m\epsilon)^{-1/2} d\epsilon$. Substituting these expressions for v^2 and dv into $f(v)dv$ with $f(v)$ given by Eq. (15-5), we obtain an expression for $g(\epsilon)d\epsilon$:

$$g(\epsilon)d\epsilon = 2\pi \left(\frac{1}{\pi k_B T} \right)^{3/2} \epsilon^{1/2} e^{-\epsilon/k_B T} d\epsilon \quad (15-11)$$

where $k_B = R/N_A$ is Boltzmann's constant. The function $g(\epsilon)d\epsilon$ represents the fraction of molecules with energies in the range between ϵ and $\epsilon + d\epsilon$. We can then

calculate the average energy of a molecule to be

$$\langle \epsilon \rangle = \int_0^\infty \epsilon g(\epsilon) d\epsilon = \frac{3}{2} k_B T = \frac{3RT}{2N_A} \quad (15-12)$$

This expression implies that the average molecular energy is greater for higher temperatures and is independent of the molecular mass.

It is useful to connect this result for the average molecular energy to thermodynamics. For an ideal monatomic gas, neglecting electronic degrees of freedom, the energy of the system arises from only translational energy. Under these assumptions, the average energy E of a system comprised of n moles of a monatomic gas is the product of the number of molecules, nN_A , and the average energy of one molecule:

$$E = nN_A \langle \epsilon \rangle = \frac{3}{2} nRT \quad (15-13)$$

This result is an example of the equipartition of energy principle, which states that all quadratic contributions to the total energy have the same mean value per mole of $(1/2) RT$. For an ideal monatomic gas, there are three quadratic terms in the three-dimensional translational energy expression, and we obtain an average molar energy of $(3/2) RT$.

15.3 BOLTZMANN DISTRIBUTION

To understand how to utilize the energy levels of molecules to calculate thermodynamic properties, we need to consider the distribution of molecular states. Some of the basic concepts from the previous section for describing the distribution of molecular speeds will also be useful for describing the distribution of molecular states. In this section, we derive the Boltzmann distribution for independent particles, neglecting the interactions between molecules. The Boltzmann distribution relates the populations of the molecular states to the energies of these states. In later sections, we will show how these general concepts can be extended to more realistic systems that include intermolecular interactions.

Consider a closed system of N noninteracting, independent molecules. Because we are neglecting interactions between molecules, the total energy of the system is simply the sum of the energies of the individual molecules. Each molecule can exist in states with energies E_0, E_1, E_2 , and so forth. For simplicity, we assume that $E_0 = 0$ and measure all other energies relative to E_0 . At any given time, there are N_0 molecules in the state with energy E_0 , N_1 molecules in the state with energy E_1 , and so forth. We define the *configuration* of the system to be the specification of the populations for all states and denote the configuration by $\{N_i\} \equiv \{N_0, N_1, N_2, \dots\}$. The *weight* W of a configuration is the number of different ways this particular configuration can be achieved. The number of ways N molecules can be arranged in N states is $N!$. However, there are $N_0!$ ways to arrange the N_0 molecules with energy E_0 , $N_1!$ ways to arrange the N_1 molecules with energy E_1 , and so forth. Therefore, the number of

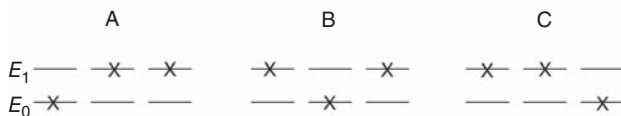


FIGURE 15-3. Schematic depiction of the three possible arrangements for a two-state system with three molecules, $N = 3$, and a configuration corresponding to $N_0 = 1$ and $N_1 = 2$, signifying that one molecule occupies the lower state and two molecules occupy the higher state. The energies of the two states are E_0 and E_1 , denoted by horizontal lines, and the occupied state for each molecule is indicated by an X. The three arrangements labeled A, B, and C are associated with the first, second, and third molecule, respectively, occupying the lower state. For this case, the weight given by Eq. (15-14) is $W = 3$, corresponding to the three possible arrangements shown here.

distinguishable ways of arranging the molecules to obtain the configuration $\{N_0, N_1, N_2, \dots\}$ must be reduced by $N_0!$, $N_1!$, and so forth, leading to

$$W = \frac{N!}{N_0!N_1!N_2! \dots} \quad (15-14)$$

This expression is just the number of ways of distributing N identical objects with a particular arrangement given by $\{N_i\}$. For example, consider the case for a two-state system with $N = 3$ and a configuration corresponding to $N_0 = 1$ and $N_1 = 2$. In this case, Eq. (15-14) tells us that $W = 3$, so there are three ways of achieving this configuration. Specifically, any of the three molecules can be in the state with energy E_0 , and the other two molecules will be in the state with energy E_1 , as depicted in Figure 15-3. In comparison, for the configuration corresponding to $N_0 = 3$ and $N_1 = 0$, Eq. (15-14) tells us that $W = 1$. In this case, there is only one way to achieve this configuration, namely for all three molecules to be in the state with energy E_0 .

For macroscopic systems, it has been shown that a single configuration has a much greater weight than all others and therefore dominates the properties of the system. In other words, the system is almost always found in this dominant configuration, and we can focus solely on this configuration to characterize the system. In this case, our challenge is to identify this dominant configuration. From a mathematical standpoint, we want to find the configuration that maximizes the weight W given in Eq. (15-14). However, we do not have complete flexibility in choosing this configuration. We can only consider the configurations that satisfy two important constraints. First, the total number of molecules in the system must remain fixed. We can define the total number of molecules to be

$$N = \sum_i N_i \quad (15-15)$$

where we are summing over all molecular states. According to this constraint, if we remove a molecule from one state, we need to add a molecule to another state to maintain the same number of molecules in the system. Second, the total energy of the system must remain fixed. The total energy of the system is defined as the sum of

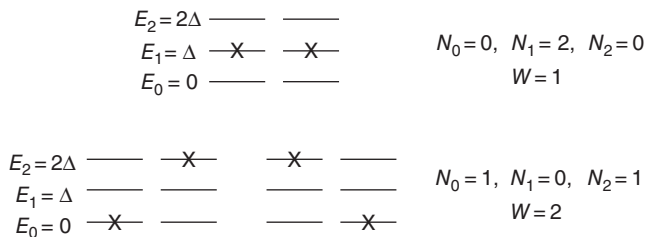


FIGURE 15-4. Schematic depiction of the allowed arrangements of two molecules over three states with energies $E_0 = 0$, $E_1 = \Delta$, and $E_2 = 2\Delta$, where the total energy of the system is constrained to be 2Δ . The upper row shows one possible configuration with $N_0 = 0$, $N_1 = 2$, and $N_2 = 0$. This configuration has weight $W = 1$, and the single possible arrangement corresponds to both molecules occupying the state with energy E_1 . The lower row shows another possible configuration with $N_0 = 1$, $N_1 = 0$, and $N_2 = 1$. This configuration has weight $W = 2$, and the two possible arrangements correspond to one molecule occupying the state with energy E_0 and the other molecule occupying the state with energy E_2 .

the energies of all molecules:

$$E = N_0 E_0 + N_1 E_1 + \cdots N_i E_i + \cdots = \sum_i N_i E_i, \quad (15-16)$$

where again the summation is over all molecular states. According to this constraint, if we move a molecule to a state of higher energy, then we must move another molecule to a state of lower energy to exactly balance the change in energy so that the total energy remains constant. A schematic picture of configurations that maintain these two constraints is depicted in Figure 15-4.

Thus, our goal is to find the configuration that maximizes the weight W while satisfying the constraints in Eqs. (15-15) and (15-16). Mathematically, it is more convenient to work with $\ln W$, the natural logarithm of W , rather than W itself. Using the expression for W given in Eq. (15-14) and the basic properties of logarithms, we obtain

$$\ln W = \ln \frac{N!}{N_0! N_1! N_2! \cdots} = \ln N! - \ln N_0! - \ln N_1! - \cdots = \ln N! - \sum_i \ln N_i! \quad (15-17)$$

This expression can be further simplified using Stirling's approximation, $\ln x! = x \ln x - x$, which is valid for large values of x , to obtain

$$\ln W = (N \ln N - N) - \sum_i (N_i \ln N_i - N_i) = N \ln N - \sum_i N_i \ln N_i \quad (15-18)$$

where the last equality arises because the sum over all N_i is N and therefore cancels the other N in the expression. Note that the maximum of $\ln W$ in the absence of constraints will occur when the second term is zero, corresponding to $N_i = 1$ for all values of i . However, this solution does not satisfy the constraints on the number of molecules

and energy given in Eqs. (15-15) and (15-16). For this purpose, the derivative of $\ln W$ with respect to each N_i is set equal to zero subject to these constraints on the possible values of N_i . The method of undetermined multipliers then allows us to find the values of N_i that maximize $\ln W$ while still satisfying these two constraints.

In solving this mathematical problem, we find that the populations in the configuration of greatest weight subject to these two constraints are given by

$$\frac{N_i}{N} = \frac{e^{-\beta E_i}}{\sum_i e^{-\beta E_i}} \quad (15-19)$$

This expression is called the *Boltzmann distribution* because Boltzmann derived it and showed that $\beta = 1/(k_B T)$, where k_B is Boltzmann's constant and T is the thermodynamic temperature. This equation shows that the most probable populations of the states of a system at thermal equilibrium are governed by the thermodynamic temperature. The left-hand side of Eq. (15-19) is $p_i = N_i/N$, which is the fraction of molecules in any state i or, equivalently, the probability that a molecule will occupy state i . The overall expression relates the populations, or occupation probabilities, of the states to the energies of the states.

In Eq. (15-19), the summation is over all states of the system, but we can rewrite this equation so that the summation is over energy levels instead of states. Summing over energy levels is convenient if several states have the same energy. If g_i states have the same energy, these states are called *degenerate* with a degeneracy of g_i . In this case, all of the states with the same energy E_i can be grouped together in a single term that is written as $g_i e^{-\beta E_i}$. The value of the total summation is exactly the same for these two types of summations, even though some terms are combined when summing over energy levels. The Boltzmann distribution involving a summation over energy levels is expressed as

$$\frac{N_i}{N} = \frac{g_i e^{-\beta E_i}}{\sum_i g_i e^{-\beta E_i}} \quad (15-20)$$

If the system is nondegenerate, so all degeneracies g_i are unity, then this expression is identical to Eq. (15-19). Even for systems with degeneracies, the populations of the energy levels calculated from Eqs. (15-19) and (15-20) are identical.

We can use the Boltzmann distribution to calculate the ratio of the number of molecules in two energy levels with energies E_i and E_j , where $E_i < E_j$, as

$$\frac{N_j}{N_i} = \frac{g_j e^{-\beta(E_j - E_i)}}{g_i} \quad (15-21)$$

From this expression, we can understand the behavior of the populations at low and high temperatures, as well as in the intermediate regime. At zero temperature, $\beta = 1/(k_B T)$ approaches infinity and $N_j/N_i = 0$, so only the ground level is populated. In this case, all of the molecules are in the lowest energy level. As the

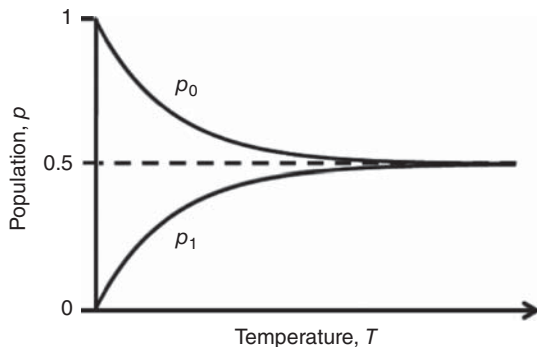


FIGURE 15-5. The populations for a nondegenerate two-state system as a function of the temperature T . The populations are $p_i = N_i/N$, as given by the Boltzmann distribution in Eq. (15-19). These populations can be interpreted as the fraction of molecules in a given state or the probability of a molecule occupying a given state. At zero temperature, only the lower state is populated, and as the temperature approaches infinity, the two states become equally populated.

temperature approaches infinity, $\beta = 1/(k_B T)$ approaches zero and $N_j/N_i \rightarrow g_j/g_i$, so the populations become independent of the energies. In this case, the molecules are distributed among the energy levels according to the degeneracies, and if there are no degeneracies, all levels are equally populated. For intermediate temperatures, a range of energy levels is populated. The populations for a nondegenerate two-state system are depicted in Figure 15-5, illustrating this behavior as a function of temperature.

The qualitative behavior of the molecular state populations at equilibrium can be analyzed for the different types of molecular energy levels. For translational and rotational energy levels, typically the spacing between levels is much less than $k_B T$ at room temperature, so many energy levels are populated. In contrast, for electronic energy levels, typically the spacing is much greater than $k_B T$ at room temperature, so only the lowest energy level is significantly populated. The vibrational energy levels are in the intermediate regime: for high-frequency vibrations, which have large energy level spacings, only the lowest energy level is significantly populated, whereas for low-frequency vibrations, several of the vibrational states are populated at room temperature.

Recall from Section 15.2 that an average quantity can be calculated as the sum of each possible value multiplied by the probability of the occurrence of that value. For the kinetic model of gases, we used this concept to calculate the average speed. Here we use this approach to calculate the average energy. The probability of a molecule occupying a state with energy E_i is $p_i = N_i/N$, which is given by the Boltzmann distribution in Eq. (15-20). Using these relations, the average energy can be computed as

$$\langle E \rangle = \sum_i p_i E_i = \sum_i \frac{N_i}{N} E_i = \frac{\sum_i g_i E_i e^{-\beta E_i}}{\sum_i g_i e^{-\beta E_i}} \quad (15-22)$$

At zero temperature, all of the molecules are in the lowest energy level, and the average energy is simply E_0 . This value is the lower bound for the average energy. As the temperature approaches infinity, all states become equally populated. In this case, the average energy becomes the average over the energies of all states: $\langle E \rangle = \frac{1}{N} \sum_i g_i E_i$, which will approach infinity for a system composed of an infinite number of molecular states. For the idealized nondegenerate two-state system depicted in Figure 15-5, the average energy is E_0 at zero temperature and approaches the average energy of the two states, $(E_0 + E_1)/2$, as the temperature approaches infinity.

15.4 MOLECULAR PARTITION FUNCTION

The denominator of the Boltzmann distribution is often defined to be the molecular partition function, given by

$$q = \sum_i g_i e^{-\beta E_i} \quad (15-23)$$

where the summation is over energy levels, and g_i denotes the degeneracies of the energy levels. The molecular partition function indicates the number of states that are thermally accessible to a collection of molecules at temperature T . The Boltzmann distribution is often expressed in terms of this quantity because of the simpler form:

$$p_i = \frac{N_i}{N} = \frac{e^{-\beta E_i}}{q} \quad (15-24)$$

The partition function can be viewed as the sum of the statistical weights for all configurations of the system, with each term being the statistical weight for that particular energy. The probability of a specified energy level is then given by the statistical weight for that energy level divided by the sum of statistical weights (Eq. (15-24)).

We can gain further insight into the physical meaning of the partition function by considering its dependence on temperature. At zero temperature, $\beta = 1/(k_B T)$ approaches infinity, and the only term that remains in the summation in Eq. (15-23) is the $E_0 = 0$ term. Thus, the partition function is simply the degeneracy of the lowest energy level, g_0 , at zero temperature. As the temperature approaches infinity, $\beta = 1/(k_B T)$ approaches zero, and all of the exponentials in the summation in Eq. (15-23) approach unity. In this limit, the partition function becomes equivalent to the number of molecular states and thus approaches infinity for a system composed of an infinite number of molecular states. For the idealized two-state system depicted in Figure 15-5, the partition function is unity at zero temperature because only the ground state is accessible and increases to two at higher temperatures because both states become accessible. If the splitting between the two energy levels is much smaller than $k_B T$, then the partition function will be nearly two at this temperature because both states are thermally accessible.

For the remainder of this section, we will derive an explicit expression for the molecular partition function of noninteracting gas molecules. In general, the energy

of a molecule is the sum of the following contributions:

$$\epsilon_i = \epsilon_i^T + \epsilon_i^R + \epsilon_i^V + \epsilon_i^E \quad (15-25)$$

where the superscripts refer to translational (T), rotational (R), vibrational (V), and electronic (E) contributions. Although the separation of these contributions is not rigorous, we assume that they are separable for simplicity. We can then substitute the molecular energy given by Eq. (15-25) into Eq. (15-23), summing over states rather than energy levels for clarity. In this way, we are able to factor the molecular partition function into four separate contributions:

$$\begin{aligned} q &= \sum_i e^{-\beta(\epsilon_i^T + \epsilon_i^R + \epsilon_i^V + \epsilon_i^E)} \\ &= \left(\sum_{i(T)} e^{-\beta\epsilon_i^T} \right) \left(\sum_{i(R)} e^{-\beta\epsilon_i^R} \right) \left(\sum_{i(V)} e^{-\beta\epsilon_i^V} \right) \left(\sum_{i(E)} e^{-\beta\epsilon_i^E} \right) \\ &= q^T q^R q^V q^E \end{aligned} \quad (15-26)$$

Here the summations labeled T, R, V, and E are over translational, rotational, vibrational, and electronic states, respectively, and the partition functions defined in the last equality correspond to these individual contributions. We will calculate each of these partition functions separately using the energy levels derived in the previous chapters on quantum mechanics.

The translational energy levels can be determined from the particle in a box model system studied in Chapter 7. The energy levels for a particle of mass m moving in a box of length L are given by

$$E_n = (n^2 h^2) / (8mL^2) \quad (15-27)$$

where the quantum number n can be any positive, nonzero integer. The lowest energy level has energy $E_1 = h^2 / (8mL^2)$, and the energies relative to this level are $E_n = (n^2 - 1)E_1$. Substituting these values into the expression for the translational partition function gives

$$q^T = \sum_{n=1}^{\infty} e^{-\beta(n^2-1)E_1}$$

Assuming that the splittings between the energy levels are much smaller than $k_B T$, the summation can be replaced by an integral and evaluated as follows:

$$q^T = \int_1^{\infty} e^{-(n^2-1)\beta E_1} dn \approx \int_0^{\infty} e^{-n^2\beta E_1} dn = (2\pi m k_B T)^{1/2} L / h \quad (15-28)$$

Here we have approximated the integral by extending the lower limit to $n=0$ and replacing n^2-1 by n^2 to produce a standard integral that is evaluated analytically in the final equality. This procedure can be extended to a particle moving in a

three-dimensional box of volume V , leading to the following expression for the translational partition function:

$$q^T = V(2\pi mk_B T)^{3/2} / h^3 \quad (15-29)$$

This partition function becomes infinity as the temperature approaches infinity because an infinite number of translational states would become accessible.

The rotational energy levels can be determined from the rigid rotor model system, equivalent to a particle on a sphere, studied in Chapter 7. The energy levels of a linear rigid rotor can be expressed as

$$E_J = hBJ(J+1) = hc\tilde{B}J(J+1) \quad (15-30)$$

Here the quantum number J can be any nonnegative integer, and the rotational constant is defined as $\tilde{B} = B/c = \hbar/(4\pi Ic)$ in units of wavenumbers (cm^{-1}), where c is the speed of light and I is the moment of inertia of the molecule. Each energy level has a degeneracy of $2J+1$. Using these quantities, we obtain the following expression for the rotational partition function of a nonsymmetric linear molecule:

$$q^R = \sum_{J=0}^{\infty} (2J+1) e^{-\beta hBJ(J+1)} \quad (15-31)$$

When the splittings between rotational energy levels are much smaller than $k_B T$, the summation can be approximated by an integral, which in turn can be solved analytically:

$$q^R = \int_0^{\infty} (2J+1) e^{-\beta hBJ(J+1)} dJ = -\frac{1}{\beta hB} \int_0^{\infty} \frac{d}{dJ} e^{-\beta hBJ(J+1)} dJ = \frac{k_B T}{hB} \quad (15-32)$$

Here the second equality involves a simple mathematical manipulation that can be confirmed by evaluating the derivative in the integral, and the third equality utilizes the property that the integral of a derivative of a function is the function itself. The greater the moment of inertia, the larger the rotational partition function because the spacing between the rotational energy levels is smaller. This property implies that more rotational states are accessible for large, heavy molecules than for small, light molecules. These expressions can be extended to nonlinear molecules, leading to more complicated equations. Moreover, for symmetrical linear molecules, the rotational partition function given earlier should be divided by a factor of two because a rotation through 180° results in an indistinguishable state of the molecule, so the number of thermally accessible states is half the number that can be occupied by a nonsymmetric linear molecule.

The vibrational energy levels can be determined from the harmonic oscillator model system studied in Chapter 7. The energy levels for a harmonic oscillator are given by

$$E_n = \left(n + \frac{1}{2}\right) h\nu \quad (15-33)$$

Here ν is the frequency of the vibrational mode, and the quantum number n can be any nonnegative integer. The lowest energy level has energy $E_0 = h\nu/2$, and the energies relative to this level are $E_n = nh\nu$. Substituting these values into the expression for the vibrational partition function gives

$$q^V = \sum_{n=0}^{\infty} e^{-\beta nh\nu} = \sum_{n=0}^{\infty} (e^{-\beta h\nu})^n = \frac{1}{1 - e^{-\beta h\nu}} \quad (15-34)$$

Here we have used the mathematical property that $e^{na} = (e^a)^n$ for the second equality and the series expression $\sum_{n=0}^{\infty} x^n = 1/(1-x)$ for the third equality. When the splitting between the vibrational energy levels, $h\nu$, is much greater than $k_B T$, the vibrational partition function is unity because only the lowest energy vibrational state is accessible. When the splitting between energy levels is much smaller than $k_B T$, this partition function becomes $q^V = k_B T/(h\nu)$ using the relation $e^{-x} \approx 1 - x$ for very small values of x . In this limit, the partition function is much greater than unity because many vibrational states are accessible.

The electronic energy levels can be calculated with the electronic structure methods described in Chapter 8. For most molecules, the splitting between the ground electronic state and the lowest excited electronic state is much greater than $k_B T$. Thus, typically the electronic partition function is unity or, if the ground state is degenerate, the electronic partition function is equal to the degeneracy g_0^E of the ground state:

$$q^E = g_0^E \quad (15-35)$$

The overall molecular partition function given in Eq. (15-26) can be expressed as the product of the translational, rotational, vibrational, and electronic contributions given in this section. Note that the expressions for these contributions are only approximate because of the various assumptions used to derive the energy levels and to derive the partition functions from these energy levels. Moreover, the expression for the overall molecular partition function given in Eq. (15-26) is valid for a single molecule in the gas phase and does not account for any type of intermolecular interactions. The next section will extend these concepts to interacting systems.

15.5 ENSEMBLES

So far our discussion of statistical mechanics has assumed that the systems are composed of noninteracting, independent molecules. In reality, molecules interact with each other, however, and we need to be able to apply statistical mechanics to interacting systems as well. The key new concept required for the extension of statistical mechanics to interacting systems is the *ensemble*. We define an ensemble to be an imaginary collection of virtual copies of a real system with certain macroscopic properties specified. Each virtual copy represents a possible state of the real system. Thus, we can view the ensemble as a probability distribution for the state of the system. For

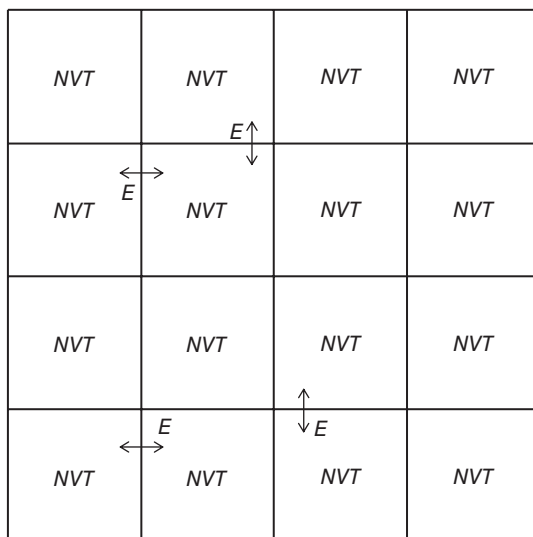


FIGURE 15-6. Schematic depiction of a canonical ensemble comprised of a large number of virtual copies of the real system. All of the virtual copies have the same number of particles N , volume V , and temperature T . The individual systems are in thermal contact and can exchange energy in the form of heat, as indicated by the double arrows labeled by E . This energy exchange occurs among all virtual copies of the system but is only shown for select pairs in this figure.

this reason, many of the statistical concepts developed in the previous sections for the distribution of molecular states are directly applicable to ensembles.

We will define three different types of ensembles, namely canonical, microcanonical, and grand canonical, where each ensemble is defined by the specification of three macroscopic properties. In a canonical ensemble, all of the virtual copies have the same number of particles N , volume V , and temperature T . Figure 15-6 depicts a physical representation of a canonical ensemble, where this type of system is replicated many times. In this representation, all of the individual closed systems are in thermal contact with each other to maintain the constant temperature T . Therefore, energy can be transferred between the virtual copies in the form of heat, and the energy of each individual system is not constant. However, the total energy of all of the virtual copies combined is fixed because the overall ensemble is an isolated system.

In a microcanonical ensemble, all of the virtual copies have the same number of particles N , volume V , and energy E . In comparison to the canonical ensemble, the energy rather than the temperature of each individual system is fixed. Because all virtual copies must have the same energy, the individual systems in a microcanonical ensemble are isolated from each other and are not in thermal contact. As a result, the temperature within the individual systems is not held constant and can vary among the different virtual copies.

In a grand canonical ensemble, all of the virtual copies have the same volume V , temperature T , and chemical potential μ . In this ensemble, the number of particles

within each individual system is no longer fixed, and the composition within each copy can fluctuate. Thus, the individual systems within a grand canonical ensemble are open, rather than closed, and molecules can pass between the individual systems. As in the canonical ensemble, the individual systems are in thermal contact with each other and can exchange energy in the form of heat, so the energy of each individual system is not fixed. The total energy and the total number of particles for all of the virtual copies combined are fixed because the overall ensemble is an isolated system.

Analogous to the discussion of molecular states in Section 15.3, we consider the various possible configurations of the ensemble. For a canonical ensemble, the configuration is defined in terms of the energies of the virtual copies within the ensemble. We assume that the total number of virtual copies in the ensemble is $\tilde{N}$, and the number of virtual copies with energy E_i is $\tilde{N}_i$. Here we use tildes to distinguish the number of virtual copies from the number of molecules discussed previously in this chapter. We can define the weights of these configurations by Eq. (15-14), simply replacing N and N_i with $\tilde{N}$ and $\tilde{N}_i$, respectively. As the number of virtual copies approaches infinity, we find that a single configuration is most probable by a large margin and therefore dominates the properties of the system. To identify this dominant configuration, we calculate the configuration of the greatest weight, subject to the constraints that the total number of virtual copies within the ensemble and the total energy of the ensemble are fixed. This procedure leads to the *canonical distribution*, which is analogous to the Boltzmann distribution given in Eq. (15-19):

$$\frac{\tilde{N}_i}{\tilde{N}} = \frac{e^{-\beta E_i}}{\sum_i e^{-\beta E_i}} \quad (15-36)$$

Analogous to the definition of the molecular partition function q in Eq. (15-23), we define the canonical partition function Q to be the denominator of this equation:

$$Q = \sum_i e^{-\beta E_i} \quad (15-37)$$

The canonical partition function contains the thermodynamic information of the system and can be used to calculate thermodynamic quantities, such as the internal energy, entropy, enthalpy, pressure, Helmholtz energy, and Gibbs energy. The next section discusses the relation between entropy and the canonical partition function.

We can also relate the canonical partition function to the molecular partition function for the special case of noninteracting, independent molecules. In this case, the total energy of the system is simply the sum of the energies of the molecules, and the canonical partition function factorizes into a product of molecular partition functions. A slight complication arises for indistinguishable molecules, which are defined as identical, noninteracting molecules, because many of the states counted in the partition function are not distinguishable. Taking this into account, the canonical partition function can be expressed as follows for a system comprised of N noninteracting,

independent molecules:

$$\begin{aligned} Q &= q^N && \text{for distinguishable independent molecules} \\ Q &= q^N/N! && \text{for indistinguishable independent molecules} \end{aligned} \quad (15-38)$$

It is important to realize that the canonical partition function is not limited to this special case but also applies to systems with interacting molecules. For the more general case of interacting systems, the expressions in Eq. (15-38) are no longer valid.

15.6 STATISTICAL ENTROPY

In the earlier chapters, we described entropy as a measure of the disorder of the system. Given this background, we might expect that entropy is related to the number of ways of distributing molecules in a system among the energy levels or, alternatively, the number of ways of distributing individual systems within a canonical ensemble. Boltzmann showed that the entropy is related to W , defined here as the statistical weight of the most probable configuration of the system, as follows:

$$S = k_B \ln W \quad (15-39)$$

This expression, which is known as the Boltzmann formula for the entropy, provides a molecular interpretation of the entropy. (This equation was also briefly discussed in Chapter 2, Eq. (2-14).)

Moreover, the temperature dependence of this statistical entropy is the same as that of the thermodynamic entropy discussed in early chapters. As the temperature decreases, the weight of the most probable configuration decreases because fewer configurations are consistent with the total energy. In other words, fewer energy levels are accessible at lower temperatures, thereby allowing a smaller number of arrangements among the energy levels. The Boltzmann formula in Eq. (15-39) tells us that the entropy decreases as the weight of the most probable configuration decreases, implying that the entropy decreases as the temperature decreases. In the limit of zero temperature, all of the molecules will be in the lowest energy state, and the most probable configuration has a weight of unity, so $W = 1$. In this limit, the statistical entropy defined in Eq. (15-39) will be zero, consistent with the third law of thermodynamics, which states that the entropy of a perfect crystal approaches zero as the temperature approaches zero.

We can also relate the entropy to the molecular or canonical partition function. If we substitute the expression for $\ln W$ given in Eq. (15-18) into Eq. (15-39), we obtain

$$S(T) = k_B \left(N \ln N - \sum_i N_i \ln N_i \right) = k_B \sum_i N_i (\ln N - \ln N_i) = -k_B \sum_i N_i \ln \frac{N_i}{N} \quad (15-40)$$

Here the second equality used Eq. (15-15) for the simplification, and the third equality used basic properties of logarithms. Utilizing the Boltzmann probability, performing

additional mathematical manipulations, and generalizing to the canonical partition function, we obtain

$$S(T) = \frac{E(T) - E(0)}{T} + k_B \ln Q \quad (15-41)$$

Here $E(T)$ is the internal energy of the system and is related to the total energy of the system defined in Eq. (15-16) by $E(T) = E(0) + E$ because all energies E_i were defined to be measured relative to the lowest energy E_0 . Moreover, the internal energy can be expressed in terms of the canonical partition function as

$$E(T) = E(0) - \left(\frac{\partial \ln Q}{\partial \beta} \right)_V \quad (15-42)$$

The derivations of these equations are beyond the scope of this book, but they illustrate that thermodynamic quantities can be expressed in terms of the partition function.

15.7 HELIX-COIL TRANSITION

As an example of the application of statistical mechanics to a biologically relevant problem, the transition between helical and coil structures in proteins is considered. We have previously discussed the structure of the α -helix, which is very prevalent in proteins (Figure 3-4). When a protein is denatured, the helical structure is destroyed and the protein can be considered to be randomly coiled. This helix-to-coil transition can be mimicked in synthetic polypeptides. For example, polyglutamic acid exists as an α -helix at low pH. When the pH is raised, the carboxyl groups on the side chains ionize, and the repulsion between the negative charges disrupt the helical structure. At a sufficiently high pH, polyglutamic acid becomes a random coil. This transition from α -helix to random coil occurs over a very narrow range of pH, as illustrated in Figure 15-7.

The helix-coil transition is an example of a cooperative process. The dominant features of the helical structure are the hydrogen bonds between peptide linkages that are four residues apart. To convert a turn in the helix to a coil, three hydrogen bonds must be broken, or conversely, to form a helical turn from a coil structure, three hydrogen bonds must be formed. This is the *initiation* process. After initiation, *propagation* of helical breakdown or formation involves only a single hydrogen bond. Simply stated, initiation is more difficult than propagation. This is the essence of a cooperative process, namely the process becomes easier after initiation. The bigger the difference between the ease of ensuing steps relative to initiation, the higher the degree of cooperativity. For the helix-coil transition illustrated in Figure 15-7, the cooperativity is shown to be thermodynamic in nature. However, as it turns out for this case, the cooperativity is also found in the kinetics of the transition.

A theoretical description of the helix-coil transition can be derived using the concepts of statistical mechanics that have been discussed in this chapter. This requires

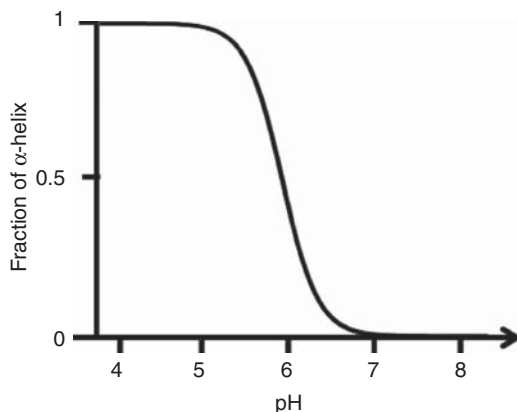
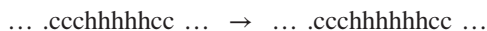


FIGURE 15-7. Schematic representation of the helix-to-coil transition in polyglutamic acid as a function of pH. Derived from data in P. Doty, *J. Polym. Sci* **23**, 815 (1957).

the development of the partition function for a polymer of amino acids. The polypeptide is assumed to be a linear chain with each amino acid designated as being in a helical or coil configuration, denoted h or c, respectively. Thus, we might have a chain such as ccchhhhhcccchh. To construct the partition function, we arbitrarily assign a statistical weight of 1 to a coil residue. This is equivalent to setting the energy to zero for the system in the coil configuration. Next we assign a statistical weight of s to a helical residue. Finally, recognizing that initiating a helix following a coil residue is more difficult than adding a helix after another helix residue, we assign a statistical weight of σs to a helical residue that initiates a helix. In other words, for an h following an h, the statistical weight is s , whereas for an h following a c, the statistical weight is σs . Cooperativity is introduced into the system by stipulating that σ is much less than 1. Using the rules outlined so far, the statistical weight of the configuration ccchhhhhcccchh is $\sigma^2 s^7$.

The statistical weights can also be viewed in another way. If we consider the addition of a helical residue to an existing helical segment such as



the equilibrium constant for this reaction is s . This is the *propagation* step in helix formation. For the addition of a helical residue into a coil sequence such as



the equilibrium constant is σs . This is the *initiation* step in helix formation. This description of the statistical weights automatically creates a temperature dependence for the statistical weights. For example, $s = \exp(-\Delta G^\circ/RT)$, where ΔG° is the standard Gibbs energy for adding a helix residue after another helix residue. In practice, σ is assumed to be a constant, independent of temperature, but this is not necessary.

The model developed thus far assumes that interactions occur only between adjacent sites, that is, no long-range interactions exist between amino acid residues. This simple but elegant model for introducing cooperativity into a linear chain containing only two elements is the essence of the Ising model developed by Ernst Ising in 1920 when he was working with Wilhelm Lenz. They used this model to describe the alignment of nuclear spins in ferromagnetism.

Unfortunately the helix-coil transition does not quite fit into the Ising model, as we know that some long-range interactions exist. As described in Chapter 9, the helix is stabilized by a hydrogen bond from a CO in residue i to an NH four residues away. Thus, if we start at the amino end of a polypeptide, the first residue hydrogen bonds to the fifth residue and so on down the polypeptide chain. However, when we get to the end of the chain, we can see that the fourth amino acid from the end cannot be in a helical conformation—it has nothing with which to hydrogen bond. Thus, the four amino acids at the end must always be in a c conformation: ...cccc. This must be taken into account when determining all of the possible configurations of the system. This problem can be easily addressed because the cccc structure persists in all configurations of the polypeptide chain.

The second related issue is more subtle. To break a helical turn requires a minimum of three hydrogen bonds to be broken so that we must have a minimum of three coil residues to break a string of helical residues. This means that a configuration such as ...hhchh... is not allowed. To take this issue into account, a statistical weight of zero is assigned to configurations in which one or two c's are interposed between two h's. Assigning a statistical weight of zero is equivalent to assuming that these configurations cannot exist. This is a rather extreme solution to this problem, as one might envision one or two hydrogen bonds being broken to give rise to states other than strictly h or c.

Using the rules that we have discussed, we can construct the partition function from the sum of statistical weights for all possible configurations of the polypeptide chain. The average number of helix residues, $\langle n_h \rangle$, can be derived analogously to the average energy, as described earlier (Eq. (15-22)). In this case, we sum over all possible configurations of the polypeptide chain, and each term is the product of the number of h's and the statistical weight for that configuration, divided by the partition function. Each nonzero statistical weight has the general formula $s^n \sigma^m$ where n is the number of helical residues and m is the number of helix initiation sites for that specific configuration. Therefore, the derivative of each statistical weight with respect to s multiplied by s is the number of helical residues in that specific configuration multiplied by the statistical weight. If this operation is performed for all configurations, and the resulting sum is divided by the partition function Q , the average number of helical residues, $\langle n_h \rangle$, is obtained. Thus, in mathematical notation, the average number of helical residues is $\langle n_h \rangle = s(dQ/ds)/Q = (d \ln Q / d \ln s)$. The fraction of the polypeptide present as a helix, f_h , is the average number of helix residues divided by the total number of residues that can be in the helix configuration:

$$f_h = \frac{\langle n_h \rangle}{n - 4}$$

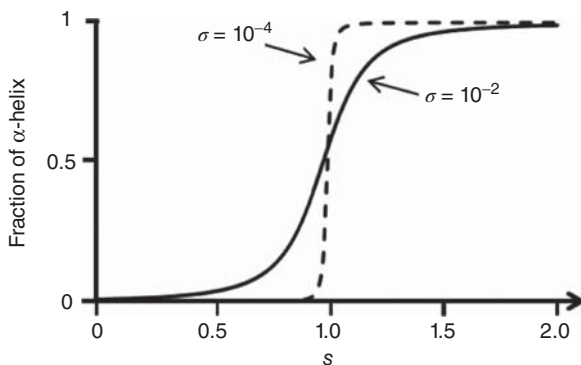


FIGURE 15-8. Schematic representation of the helix-coil transition for two different values of σ , the factor by which the statistical weight is smaller for initiating a helix after a coil residue than for propagating an existing helix. The abscissa is the equilibrium constant for propagating a helix, which corresponds to adding a helix residue after another helix residue. This plot was derived from the data in B. H. Zimm and J.K. Bragg, *J. Chem. Phys.*, **31**, 526 (1959).

Here n is the number of residues in the polypeptide chain, and the four in the denominator arises because the last four residues are not counted as they can never be an h residue. A plot of the fraction of helix versus s for large n and two values of σ is shown in Figure 15-8.

The transition from coil to helix is attributed to a small change in s . In the case of polyglutamic acid, this change comes about because of the ionization of carboxyl groups, but this change can also be due to temperature changes because s is temperature dependent. Moreover, the smaller the value of σ , the sharper, more cooperative the transition. This is to be expected: the more difficult the initiation, the more cooperative the transition. Also, the more cooperative the transition, the more sigmoidal the curve (Figure 15-8). This theory of helix-coil transitions is found to fit experimental data well with σ equal to about 10^{-4} for large values of n . The sharp transition from helix to coil as the temperature is raised is similar to the sharp transition from native to denatured structure observed for proteins.

This statistical mechanical analysis of the helix-to-coil transition illustrates how models for biological processes can be developed and understood through statistical mechanics. The next chapter will deal with more applications of theory to the analysis of biological problems.

REFERENCES

The topics in this chapter are discussed in considerably more depth in a number of physical chemistry textbooks, such as References 1–5.

1. I. Tinoco, Jr., K. Sauer, & J. C. Wang, J. D. Puglisi, G. Harbison, and D. Rovnyak, *Physical Chemistry: Principles and Applications to the Biological Sciences*, 5th edition, Prentice Hall, Englewood Cliffs, NJ, 2013.

2. P. Atkins and J. de Paul, *Physical Chemistry*, 9th edition, W. H. Freeman and Co., New York, NY, 2010.
3. P. L. Houston, *Chemical Kinetics and Reaction Dynamics*, Dover Publications, New York, NY, 2006, first published by McGraw-Hill Companies, New York, NY, 2001.
4. D. A. McQuarrie, *Statistical Mechanics*, University Science Books, Sausalito, CA, 2000.
5. D. Chandler, *Introduction to Modern Statistical Mechanics*, Oxford University Press, New York, NY, 1987.
6. B. H. Zimm and J.K. Bragg, *J. Chem. Phys.*, 31, 526 (1959).

PROBLEMS

- 15-1.** Consider a collection of noninteracting ideal gas molecules with molar mass M at temperature T .
- a. What is the ratio of the probability of finding a molecule moving with the average speed to the probability of finding a molecule moving with two times the average speed?
 - b. As the temperature increases, how does this ratio change?
 - c. As the molar mass increases, how does this ratio change?
- 15-2.** Consider a system of N identical noninteracting molecules. The molecules can exist in three energy states with energies 0, ϵ_1 , and ϵ_2 .
- a. Derive expressions for the molecular partition function and the canonical partition function.
 - b. Derive an expression for the fraction of molecules in each energy state. What is the fraction of molecules in each energy state at very high temperatures, that is, as T approaches infinity? What is the fraction of molecules in each energy state at very low temperatures, that is, as T approaches 0?
 - c. Derive an expression for the average energy of the system. What is the average energy at very high temperatures, that is, as T approaches infinity? What is the average energy at very low temperatures, that is, as T approaches 0?
- 15-3.** Consider the molecular partition functions for two diatomic molecules, HBr and Br₂.
- a. Calculate the rotational and vibrational partition functions at 500 K for HBr, given a rotational constant of $\tilde{B} = 8.465 \text{ cm}^{-1}$ and a vibrational frequency of $\tilde{\nu} = 2649 \text{ cm}^{-1}$, where $\nu = \tilde{\nu}c$ and c is the speed of light.
 - b. Based on your calculations, how many vibrational states are accessible for HBr at this temperature? Are more or fewer rotational states accessible?
 - c. Calculate the rotational and vibrational partition functions at 500 K for Br₂, given a rotational constant of $\tilde{B} = 0.0821 \text{ cm}^{-1}$ and a vibrational frequency of $\tilde{\nu} = 325 \text{ cm}^{-1}$.
 - d. Based on your calculations, are there more rotational states accessible for HBr or Br₂? What is the physical basis for this observation?

- e. Based on your calculations, are there more vibrational states accessible for HBr or Br₂? What is the physical basis for this observation?
- 15-4.** Consider an ideal monatomic gas comprised of $N = nN_A$ molecules, where n is the number of moles and N_A is Avogadro's number.
- Using Eq. (15-41), calculate the entropy of this system. You can use the expression for the canonical partition function, $Q = q^N/N!$ given in Eq. (15-38) for indistinguishable particles. You can also use Stirling's approximation, $\ln x! = x \ln x - x$, because the number of molecules in a typical sample is large. In addition, you can use the translational partition function given in Eq. (15-29) because the only motion for an atomic gas is translational. Finally, you can use Eq. (15-13) to obtain an expression for $E = E(T) - E(0)$. The resulting expression is called the Sackur-Tetrode equation.
 - Using the ideal gas law, $PV = nRT$, rewrite this expression in terms of pressure instead of volume.
 - As the molar mass increases, does the molar entropy increase or decrease? What is the physical basis for this trend?
- 15-5.** Consider a peptide comprised of eight amino acids in the context of the helix-coil transition. Note that the helix is most likely not very stable with only eight amino acids, and we use this problem for instructive purposes only.
- Using the rules discussed earlier for the statistical weights, calculate the partition function Q , which is the sum of the statistical weights for all configurations of this system.
 - Calculate the average number of helix residues, $\langle n_h \rangle$.
 - Calculate the fraction of the polypeptide present as a helix, f_h .

Molecular Simulations

16.1 INTRODUCTION

Computer simulations serve as a bridge between theory and experiment. In particular, simulations can provide the numerically exact results for a specific model. The validity of the model can be tested by comparing the simulated results to experimental data for the system that is being modeled. Moreover, theories can be tested by comparing the simulated results to theoretical predictions for this model. The combination of theories, models, and simulations assist in the interpretation of experimental results and provide predictions that are experimentally testable. Furthermore, computer simulations serve as a bridge between microscopic and macroscopic properties. Specifically, simulations provide microscopic details on the atomic level, while also providing macroscopic thermodynamic properties. Finally, computer simulations can provide information that is not readily available from experiments, such as measurements under extreme conditions of temperature and pressure that are experimentally prohibitive, details of molecular motion and structure that are difficult or impossible to study experimentally, and molecular events that are too fast or slow to study experimentally. The importance of molecular simulations was recognized by the 2013 Nobel Prize in Chemistry, which was awarded to Martin Karplus, Michael Levitt, and Arieh Warshel for their pioneering work on computer modeling of complex chemical systems.

The complete mathematical description of a system requires the solution of the Schrödinger equation for all electrons and all nuclei, but these types of simulations are not computationally practical for large biological systems or even most molecular systems. Instead we rely on the Born–Oppenheimer approximation, which decouples the motion of the electrons from the nuclei because the electrons move much faster than the nuclei due to the difference in mass. As discussed in Chapter 8, according to the Born–Oppenheimer approximation, the electrons are assumed to move in the field of fixed nuclei, and the nuclei are assumed to move in the average field of the electrons. In other words, the nuclei move on a potential energy surface obtained by solving the electronic Schrödinger equation for each nuclear configuration. To describe the nuclear motion, we need a complete description of the potential energy

surface on which the nuclei move. Rather than generating this potential energy surface by repeatedly solving the electronic Schrödinger equation, which is computationally expensive, the potential energy surface is often represented by a *molecular mechanical force field*, which can be viewed as an analytical functional form that approximates the surface. The quantum mechanical description of the nuclear motion on the potential energy surface is called *quantum dynamics*, and the classical description of the nuclear motion is called *classical molecular dynamics*. If the time evolution of the nuclei is not of interest, but rather only average properties are desired, then the system can also be described with the *Monte Carlo* approach.

In this chapter, we describe the tools that enable computer simulations of biological systems and illustrate the type of information that can be obtained from such simulations. First, we explain how molecular mechanical force fields are used to describe the potential energy surfaces, and then we summarize the methods that have been developed for large-scale simulations, such as periodic boundary conditions and treatment of long-range electrostatics. After explaining how the system is modeled, we discuss the classical molecular dynamics and Monte Carlo approaches. Subsequently, we review hybrid approaches that mix quantum mechanical and molecular mechanical methods to allow the description of chemical bond breaking and forming, as well as approaches for the calculation of Gibbs or Helmholtz energies in biological systems. The power of these types of approaches is illustrated by a discussion of computer simulations of the hydride transfer reaction catalyzed by the enzyme dihydrofolate reductase (DHFR) in the final section of this chapter.

16.2 POTENTIAL ENERGY SURFACES

A molecular mechanical force field consists of analytical functional forms describing the interactions in a system and parameters within these functional forms. The force field is usually fit to structures and relative energies obtained from experimental data or high-level electronic structure calculations for a set of molecules. In some cases, the force field may also be fit to bulk properties. Thus, the essential quantum mechanical effects, such as covalent bonding, are included empirically in the force field. In all-atom force fields, each atom is assigned a spatial coordinate, and the potential energy is defined to be a sum of interaction energies among the atoms. In some force fields, often called *coarse grained*, groups of atoms are combined into a single site, and the interaction energies are defined in terms of these sites rather than atoms to save computational time. For simplicity, this discussion will focus on all-atom force fields, although the basic concepts apply to these alternative descriptions as well.

Typically the total potential energy is expressed as

$$U = U_{\text{bonded}} + U_{\text{el}} + U_{\text{vdW}} \quad (16-1)$$

where U_{bonded} denotes bonding or intramolecular interactions, U_{el} denotes electrostatic or Coulombic interactions, and U_{vdW} denotes the van der Waals interactions, which are defined as the sum of the London attraction and van der Waals repulsion

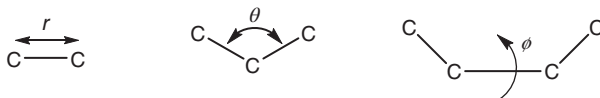


FIGURE 16-1. Stretching of the bond length r , bending of the bond angle θ , and rotation around bonds described by the torsional angle ϕ .

interactions. The main bonding interactions are bond stretching, bond angle bending, and rotation around bonds, as depicted in Figure 16-1. In some cases, additional terms are included to describe out-of-plane deformations for planar systems or cross terms to describe couplings between the various types of bonding interactions. Moreover, sometimes additional terms are included to account explicitly for hydrogen-bonding interactions. In the rest of this section, we discuss the interaction terms in Eq. (16-1) and their significance for the simulation of biological molecules.

Bond stretching can be described as if the two atoms were connected by a spring. The energy required to compress or stretch the spring is given by the harmonic oscillator potential described in Chapter 7:

$$U_b = \frac{1}{2}k_b(r - r_{\text{eq}})^2 \quad (16-2)$$

where r_{eq} is the equilibrium bond length and k_b is the force constant of the spring. Typical values of the force constants are $k_b \approx 600 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$ for single bonds and $k_b \approx 1180 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$ for double bonds. Based on these values, changing the bond length by $0.1 \text{ \AA}$ requires approximately 3 kcal/mol of energy for a single bond and approximately 6 kcal/mol of energy for a double bond. Thus, a significant amount of energy is required to stretch or compress the bond away from its equilibrium bond length. Moreover, the energy required for bond stretching or compression is greater for double bonds than that for single bonds. In other words, double bonds are stiffer than single bonds, as expected. The bond-stretching interaction terms for a single and a double bond are depicted in Figure 16-2.

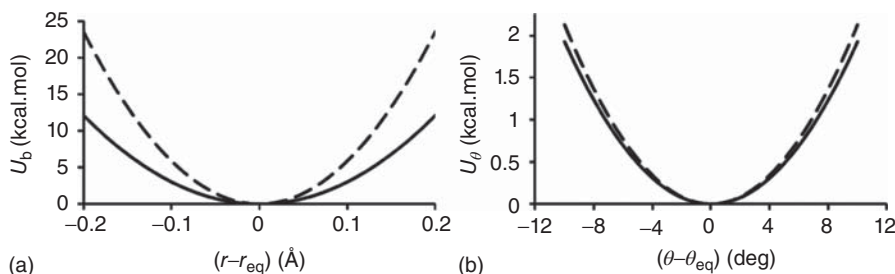


FIGURE 16-2. The harmonic potentials describing bond stretching, given by Eq. (16-2), are shown in (a) for a single (solid) and double (dashed) C–C bond. The harmonic potentials describing bond angle bending, given by Eq. (16-3), are shown in (b) for a tetrahedral (solid) and trigonal (dashed) CCC angle. *Source:* Courtesy of American Chemical Society, Biochemistry, Dec 2011.

The bending of a bond angle is described by the same type of harmonic potential:

$$U_{\theta} = \frac{1}{2}k_{\theta}(\theta - \theta_{\text{eq}})^2 \quad (16-3)$$

where θ_{eq} is the equilibrium bond angle and k_{θ} is the associated force constant. A typical value for the force constant associated with angle bending in a tetrahedral bond is $k_{\theta} \approx 0.04 \text{ kcal mol}^{-1} \text{ degree}^{-2}$. Thus, changing the bond angle by 10° requires approximately 2 kcal/mol of energy. The energy required for bond bending is slightly greater for a trigonal bond than for a tetrahedral bond, as shown in Figure 16-2.

The rotation around bonds is described in terms of torsion angles, as defined in Figure 16-1. These interaction energies are represented by periodic functions with maxima and minima at certain orientations that depend on the nature of the bonds. For example, the torsional interaction energy for rotation about a C–C single bond is

$$U_{\phi} = \frac{V_3}{2}[1 + \cos(3\phi)] \quad (16-4)$$

where V_3 determines the barrier to rotation because $U = V_3$ for $\phi = 0$. As shown in Figure 16-3, this interaction energy has three minima as the torsion angle varies from 0° to 360° for two tetrahedral carbons. For the specific case shown in Figure 16-3, the minima correspond to the two *gauche* conformations with $\phi = 60^{\circ}$ and 300° and the single *trans* conformation with $\phi = 180^{\circ}$. For a C–C double bond, the torsional interaction energy is

$$U_{\phi} = \frac{V_2}{2}[1 + \cos(2\phi - 180)] \quad (16-5)$$

In this case, the interaction energy has two minima at 0° and 180° corresponding to the planar molecule in the *cis* and *trans* conformation for the case shown in Figure 16-3. Typically $V_2 > V_3$ because more energy is required to rotate around a double bond than a single bond due to π bonding within the double bond.

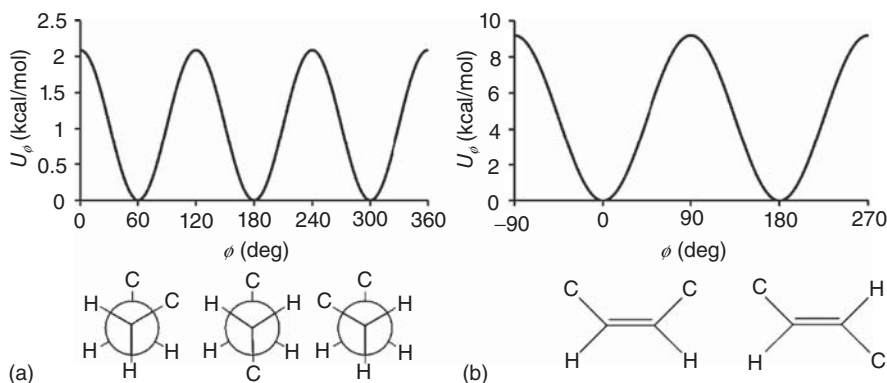


FIGURE 16-3. Torsional interaction energy for rotation about a C–C single bond (a) and a C–C double bond (b). *Source:* Courtesy of American Chemical Society, Biochemistry, Dec 2011.

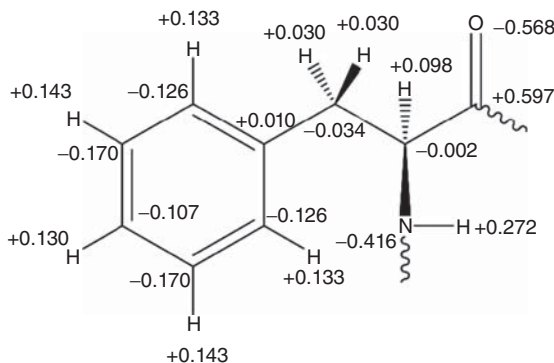


FIGURE 16-4. Example of partial atomic charges for the amino acid phenylalanine. The partial charges for each amino acid will sum to an integer charge that is consistent with the protonation state of the residue. The wavy lines represent the location of peptide bonds between the previous and subsequent amino acid residues in the protein.

In addition to these bonding interactions, nonbonding interactions between pairs of atoms also play an important role in force fields. Typically atoms that are involved in bonding interactions with each other are excluded from these nonbonding interactions because the local electronic structure is represented by the bonding terms. Nonbonding interactions include both electrostatic and van der Waals interactions, which are discussed in more detail in the remainder of this section.

The electrostatic interactions are described in terms of partial charges assigned to each atom. The partial charges can be obtained from quantum chemistry calculations on molecular systems. As discussed in Chapter 8, quantum chemistry calculations provide information about the electron density, which varies over the molecular system. Analysis tools have been developed to assign partial charges to each atom in a way that reproduces the quantum mechanical electrostatic potential while maintaining the correct overall charge of the molecule. An example of the partial atomic charges for the amino acid phenylalanine is given in Figure 16-4. These partial atomic charges can also be used to calculate the dipole moment of the molecule, which provides an indication of its polarity. Interaction terms that explicitly involve dipole moments and higher-order moments are typically not included in force fields, but rather these effects are incorporated via the interaction energies between atomic point charges.

The electrostatic interaction energy for two charges q_1 and q_2 separated by a distance r is given by

$$U_{\text{el}}(r) = 332 \frac{q_1 q_2}{r} \text{ kcal/mol} \quad (16-6)$$

The constant of 332 converts the energy to kcal/mol when the charges are given in units of electronic charge and the distance is given in units of Angstroms. This interaction energy is attractive for pairs of atoms with charges of opposite sign and repulsive for pairs of atoms with charges of the same sign, as shown in Figure 16-5. The total electrostatic interaction energy is obtained by summing over all pairs of atoms in the system except for those excluded because of bonding interactions.

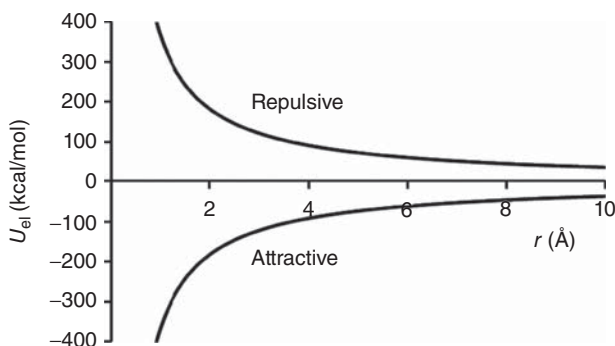


FIGURE 16-5. Electrostatic interaction energy for two charges with the same sign (repulsive interaction) and for two charges with the opposite sign (attractive interaction).

When the two charges are separated by a medium, such as solvent, with dielectric constant D , the electrostatic interaction in Eq. (16-6) is divided by this dielectric constant. The dielectric constant is unity in vacuum, approximately 80 for water, and approximately 2 for nonpolar liquid hydrocarbons. In all-atom force fields, typically the dielectric constant is assumed to be unity because the effects of the medium are included explicitly through interatomic interactions. In some cases, molecular simulations do not include explicit water molecules but instead treat the aqueous solvent as a dielectric continuum with the dielectric constant of water. Such treatments are computationally faster but are unable to describe certain phenomena, such as specific hydrogen-bonding interactions between solvent molecules and the biomolecule.

A number of issues arise in the description of electrostatic interactions. As discussed in the next section, electrostatic interactions are long range and therefore must be treated with care for large systems, particularly for periodic systems. Moreover, the treatment of the electrostatic interaction energies described earlier neglects polarization effects because the partial atomic charges are fixed during the simulation. A number of polarizable force fields have been developed to include polarization effects by allowing the partial atomic charges to fluctuate or by adding an additional charge site that is allowed to move in response to the local electric field. These polarizable force fields have been shown to be more accurate in some cases but are also more computationally expensive.

In addition to electrostatic interactions, most force fields also include the London attraction and van der Waals repulsion interaction energies. The London attraction energy is a quantum mechanical phenomenon that arises from fluctuations in the electronic distribution. These fluctuations lead to an instantaneous dipole in one atom that induces an oppositely directed dipole in a neighboring atom. This fluctuation dipole-induced dipole interaction is always attractive and becomes more negative as the distance between the two atoms decreases. The form of this interaction is often given as

$$U_{\text{att}}(r) = -\frac{B_{ij}}{r_{ij}^6} \quad (16-7)$$

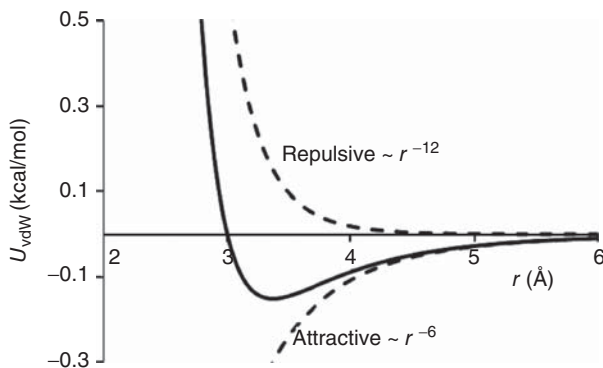


FIGURE 16-6. The dashed lines depict the London attraction interaction energy (Eq. (16-7)) and van der Waals repulsive interaction energy (Eq. (16-8)). The solid line depicts the London–van der Waals interaction energy $U_{\text{vdW}}(r)$ (Eq. (16-9) or Eq. (16-10)). Using the form given in Eq. (16-10), the well depth is ϵ_{ij} , and the position of the minimum is at an interatomic separation of $(2)^{1/6}\sigma_{ij}$.

Here B_{ij} is a positive constant that depends on the polarizability, which is a measure of the ease of delocalization of the electrons, of the groups involved. Figure 16-6 depicts this attractive interaction with a dashed line.

The van der Waals repulsion energy reflects the steric repulsions between all atoms or molecules at short distances. The origin of this interaction is also quantum mechanical in that two electrons cannot occupy a single electronic orbital, as discussed in Chapter 8. Several different simple models have been used to describe the van der Waals repulsion. In the hard-sphere model, the interaction between two atoms is zero until they come into contact, at which point the interaction energy goes to infinity. This model can be understood by envisioning two billiard balls that do not sense each other until they collide, at which point they experience a hard collision and are deflected in different directions. A more realistic model is given as

$$U_{\text{rep}}(r) = \frac{A_{ij}}{r_{ij}^{12}} \quad (16-8)$$

where A_{ij} is a positive constant that depends on the effective sizes of both atoms. Figure 16-6 depicts this repulsive interaction with a dashed line.

The van der Waals interaction energy is defined to be the sum of the London attraction and van der Waals repulsion interaction energies. Using Eqs. (16-7) and (16-8), the van der Waals interaction energy for two atoms i and j separated by distance r_{ij} is given by

$$U_{\text{vdW}}(r_{ij}) = \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \quad (16-9)$$

This interaction energy is also called the *Lennard-Jones 12–6 potential* because of the powers of the distance in the denominators of the two terms. Often it is written in

a slightly different form as

$$U_{\text{vdW}}(r_{ij}) = 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] \quad (16-10)$$

The solid line in Figure 16-6 depicts this 12–6 interaction potential, which is the sum of the attractive London interaction and the repulsive van der Waals interaction. At short distances, the repulsive term dominates because it is described by a higher power of $1/r$, resulting in a strongly repulsive region of positive potential at small distances. At larger distances, the attractive interaction dominates, resulting in an attractive region with negative interaction energy as the potential approaches zero. At intermediate distances, the attractive and repulsive interactions counteract each other, and the minimum of the curve corresponds to a net zero force due to this balance.

For the van der Waals interaction given by Eq. (16-10) and depicted in Figure 16-6, the well depth is ϵ_{ij} , and the position of the minimum is at an interatomic separation of $(2)^{1/6}\sigma_{ij}$. In practice, each atom i is assigned van der Waals parameters ϵ_i and σ_i , and the pairwise interactions are calculated using prescribed combining rules. For example, the values used to calculate the interaction energy for atoms i and j are combined as $(\sigma_i + \sigma_j)/2$ and $\sqrt{\epsilon_i\epsilon_j}$ in some force fields. For atoms such as carbon, oxygen, and nitrogen, typically the values of σ are $\sim 3.0 - 3.5$ Å, and the values of ϵ are $\sim 0.1 - 0.2$ kcal/mol.

In simulations of biological systems based on molecular mechanical force fields, the user must input the topology of the biomolecule to indicate the atom types and the bonding pattern that identifies which atoms are bonded to each other. Depending on the topology, the bonding interactions comprised of bond stretching, angle bending, and torsional angle rotation are defined for the system. Typically the topology is fixed, so bonds are not allowed to break and form. Each atom is also assigned a partial charge to describe the electrostatic interactions, as well as van der Waals parameters. All of this information is used to generate the complete potential energy surface, which depends on the nuclear coordinates, as given in Eq. (16-1).

16.3 MOLECULAR MECHANICS AND DOCKING

Molecular mechanical force fields can be used to identify the lowest energy conformations of biological systems through an approach that is often called *molecular mechanics*. These types of calculations can be used to perform structural analyses, to model the energetics of small conformational changes, and to examine how two molecules will interact with or bind to each other. Several of the issues discussed in Chapter 8 for calculating minima on the potential energy surfaces generated with quantum chemistry methods are also relevant here. The main difference is that here we are using molecular mechanical force fields instead of quantum chemistry methods to calculate the potential energy surface.

It is important to recall that the absolute energies are not meaningful, so the potential energies of different molecules cannot be compared directly. However, the relative

energies of different conformations of chemically identical systems can be compared. Minimization schemes can be used to locate minima on the potential energy surface. As discussed in Chapter 8, a minimum on the potential energy surface is defined to have zero forces and positive second derivatives along all coordinates. Due to the high dimensionality of biological systems, there are many local minima in which the system can become trapped, and it is difficult to determine if a particular minimum is the global minimum. Thus, often molecular dynamics or Monte Carlo simulations are performed to explore the potential energy surface more broadly.

Molecular mechanics can also be used to investigate how a protein or nucleic acid will bind a drug molecule or ligand. This procedure is often called *docking*. In this case, the focus is on the interaction energy between the two molecules. This interaction energy is calculated by subtracting the energies of the individual molecules from the energy of the total system in which the two molecules are interacting. Usually, the geometries of the individual molecules and the total system are optimized to minimize the energy. In some cases, the molecules are kept rigid, although flexibility of both the protein and the ligand can play a significant role in ligand binding. Typically, these approaches minimize the energy, rather than the Gibbs or Helmholtz energy, so entropic effects are not included, although methods have been developed to include entropic effects as well.

16.4 LARGE-SCALE SIMULATIONS

In biological simulations, the goal is to simulate a macroscopic system while including a relatively small number of atoms for computational tractability. A simple example is simulating a box filled with explicit water molecules. A slightly more complicated example is simulating a protein surrounded by hundreds or even thousands of explicit water molecules in a box. In these types of simulations, care must be taken in the treatment of the boundaries because a large fraction of atoms lie on the surface and thus experience different interaction energies than do the atoms in the bulk. Several methods have been developed to address this issue. One approach is to impose a spherical boundary while treating the outer shell of the sphere in a way that mimics the effects of bulk solvent. A variety of such methods have been developed and applied to biological systems.

Another popular approach for treating the boundaries in large simulations is utilizing periodic boundary conditions. In this approach, the system is placed in a cubic box that is replicated to form an infinite three-dimensional lattice. A two-dimensional representation of this lattice is depicted in Figure 16-7. As an atom moves in the central box, all of its images in the replicated boxes move in exactly the same way. If a molecule leaves the central box, an image enters through the opposite face. This approach is based on the assumption that the properties of a relatively small, infinitely periodic system are the same as those of a macroscopic system. In practice, the validity of this assumption should be tested by performing simulations with different box sizes. The central box can also be rectangular, octahedral, or another shape that is suitable for the biomolecule of interest.

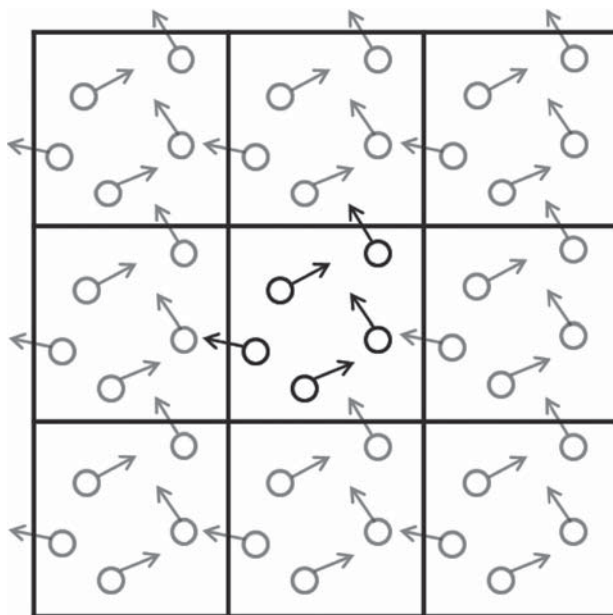


FIGURE 16-7. Schematic depiction of periodic boundary conditions in two dimensions. As a particle leaves the central box, another particle replaces it from the other side.

The calculation of the nonbonding interaction terms is computationally expensive for large systems due to the large number of pairwise interactions. For a system with N atoms, there are $N(N-1)/2$ unique pairs of atoms. To save computational time, often a spherical cutoff is applied to some of the nonbonding terms, where the interaction is set to zero at a cutoff radius r_c :

$$U(r) = 0 \text{ for } r > r_c \quad (16-11)$$

The van der Waals interactions given in Eq. (16-10) are short range in that they approach zero relatively quickly at small distances. Thus, these interactions can usually be truncated at a cutoff radius of $\sim 10\text{--}12 \text{ \AA}$. However, a sudden cutoff of nonbonded interactions can lead to discontinuities in the potential energy and the forces, leading to numerical instabilities in the molecular dynamics simulations described in the next section. To avoid these discontinuities, the potential energy is often modified slightly by the addition of terms that make the potential energy and forces zero at the cutoff radius.

Electrostatic interactions are long range in that they approach zero much more slowly and at larger distances than the short-range van der Waals interactions. As a result, using cutoffs for electrostatic interactions can lead to significant errors in simulations of large systems. A mathematically rigorous approach for summing over the interactions between ions and their periodic images is the *Ewald summation method*, named after Paul Peter Ewald, a German physicist and crystallographer.

In this method, the summation of long-range interactions is converted into two summations that can be treated as short-range interactions with cutoffs. An alternative approach is the *reaction field method*, in which the charges inside a spherical cavity are treated explicitly, while the charges outside this cavity are represented as a dielectric continuum to mimic their effects in an average way.

In Chapter 15, we discussed several different types of ensembles, where an ensemble was defined to be an imaginary collection of virtual copies of a real system with certain macroscopic properties specified. A molecular simulation is usually chosen to represent one of these types of ensembles. In a microcanonical ensemble, the number of particles, volume, and energy are held constant (constant NVE). For a box of fixed volume with a fixed number of atoms, this ensemble is generated by solving Newton's equations of motion, which conserve energy. In a canonical ensemble, the number of particles, volume, and temperature are held constant (constant NVT). For a box of fixed volume with a fixed number of atoms, this ensemble is generated by connecting the system to a heat bath that maintains a constant temperature. The energy of the system in a canonical ensemble can vary because of heat exchange between the system and the bath, but for large N usually the energy does not fluctuate significantly. Similarly, the temperature of the system in a microcanonical ensemble can vary, but for large N usually the temperature does not fluctuate widely. Another ensemble that is useful for molecular simulations is the isothermal–isobaric ensemble, where the number of particles, pressure, and temperature are fixed (constant NPT). In this ensemble, the volume of the box can change to maintain a constant pressure. This ensemble is often used to obtain the correct density of the system for a specified pressure. For large systems, the structural and thermodynamic properties tend to be similar for these commonly used ensembles.

In molecular simulations, a macroscopic property can be calculated by averaging this property over all members of the ensemble. According to the *ergodic hypothesis*, averaging a given property over all members of the ensemble is equivalent to averaging this property over time. In other words, the average over conformational space is equivalent to the average over a time-dependent trajectory. Such a time-dependent trajectory can be generated with molecular dynamics, which will be discussed in the next section. Another way to calculate properties averaged over an ensemble is Monte Carlo, which will be discussed in a later section. In addition to providing macroscopic, thermodynamically averaged properties, molecular dynamics also provides information about the time evolution and the fluctuations of the system at the atomic level.

16.5 MOLECULAR DYNAMICS

In classical molecular dynamics simulations, Newton's equations of motion are integrated for all atoms in the system. Molecular dynamics can be used to generate statistical ensembles, which in turn are used to calculate energetic, thermodynamic, structural, and dynamical properties. It can also be used to study the time evolution of chemical and biological processes. In this section, we explain the basic principles underlying molecular dynamics and discuss some of the issues that arise in

molecular dynamics simulations, such as choosing an appropriate time step, applying constraints to certain bonds, and defining and controlling the temperature of the system.

Newton's equations of motion can be written as

$$\mathbf{F}_i = m_i \mathbf{a}_i = m_i \frac{d^2 \mathbf{r}_i}{dt^2} \quad (16-12)$$

This equation states that a force $\mathbf{F}_i$ on a particle of mass m_i causes an acceleration $\mathbf{a}_i$, which is defined to be the second derivative of the position coordinate with respect to time. Here the force, acceleration, and position are vectors in three-dimensional space with x , y , and z components. Moreover, the force is the negative of the derivative of the total potential energy U of the system:

$$F_{ix} = -\frac{\partial U}{\partial r_{ix}} \quad (16-13)$$

where r_{ix} is the x -component of the position coordinate for particle i , and F_{ix} is the x -component of the force on particle i . Analogous equations can be written for the y and z components. Numerical methods are required to solve Newton's equations of motion for more than two independent particles. In these numerical methods, the positions and velocities at a given time t are used to obtain the positions and velocities at a later time $t + \Delta t$, where Δt is a small enough time step to allow the accurate solution of these equations of motion. The optimal numerical method satisfies two important properties. First, it conserves energy, where the sum of kinetic and potential energies is constant. Second, it is time reversible, where reversing the signs of all velocities results in the atoms retracing the trajectory back to their original positions and velocities. No algorithm conserves energy perfectly or is exactly time reversible over long times, but the goal is to design an algorithm that maintains these properties as well as possible.

Several algorithms have been developed to integrate Newton's equations of motion in molecular dynamics simulations. Here we will present the *velocity Verlet algorithm*, keeping in mind that many alternative algorithms are also available. This algorithm is an extension of the original Verlet algorithm named after Loup Verlet, a French physicist who pioneered molecular dynamics simulations. In this algorithm, we start with the positions $\mathbf{r}(t)$ and velocities $\mathbf{v}(t)$ and calculate $\mathbf{a}(t)$ from Eq. (16-12) using the force obtained from the potential energy surface via Eq. (16-13). Then the positions and velocities at time $t + \Delta t$ are calculated using the following equations:

$$\mathbf{r}(t + \Delta t) = \mathbf{r}(t) + \Delta t \mathbf{v}(t) + \frac{1}{2} [\Delta t]^2 \mathbf{a}(t) \quad (16-14)$$

$$\mathbf{v}(t + \Delta t) = \mathbf{v}(t) + \frac{1}{2} \Delta t [\mathbf{a}(t) + \mathbf{a}(t + \Delta t)] \quad (16-15)$$

Note that the acceleration at time $t + \Delta t$ as well as at time t is required to propagate the velocity in this algorithm. This process of numerically propagating the positions

and velocities is repeated many times to generate a molecular dynamics trajectory. The velocity Verlet algorithm is numerically stable, simple to program, and time reversible. It also conserves energy well even with relatively long time steps.

An approach such as the velocity Verlet algorithm can be used to propagate the positions and velocities of all atoms in a large biological system. It is important to choose a suitable time step for this numerical integration. Using a large time step will decrease the computational time, but if the time step is too large, the numerical integration will become unstable and inaccurate. In practice, the stability of the numerical integration can be tested by analyzing energy conservation and time reversibility. Many algorithms, including the velocity Verlet algorithm, assume that the accelerations and velocities are constant over a given time step. Thus, the time step is limited by the highest frequency motion that must be simulated. In large-scale simulations, often the bonds involving hydrogen, such as the C–H bonds, are associated with the highest frequency motions and therefore are constrained, that is, fixed to their equilibrium bond lengths. Freezing these high-frequency motions with constraints allows a larger time step to be used and therefore speeds up the computation. A typical time step for all-atom classical molecular dynamics simulations is $\sim 1\text{--}2$ fs.

Currently, molecular dynamics trajectories are often propagated for tens or hundreds of nanoseconds. More efficient algorithms and faster computers have enabled researchers to propagate molecular dynamics trajectories on the microsecond and even longer time scales. Coarse graining approaches, where the system is described in terms of sites comprised of many atoms, have allowed even longer molecular dynamics trajectories. However, these types of coarse graining approaches do not retain the atomic-level information that all-atom force fields provide.

The temperature for a molecular dynamics simulation is proportional to the kinetic energy of the system, which can be expressed in terms of the atomic velocities. This relation can be understood in the context of the Maxwell–Boltzmann distribution of velocities discussed in Chapter 15. In Eq. (15-3), we expressed the mean-square speed, $\langle v^2 \rangle$, in terms of the temperature. Rewriting this expression in terms of the mass m of a single particle or atom,

$$\langle v^2 \rangle = \frac{3k_{\text{B}}T}{m} \quad (16-16)$$

The kinetic energy of a system comprised of N particles with masses m_i and speeds v_i is

$$K = \frac{1}{2} \sum_{i=1}^N m_i v_i^2 \quad (16-17)$$

where the summation is over all particles. Combining equations (16-16) and (16-17), the average kinetic energy is

$$\langle K \rangle = 3Nk_{\text{B}}T/2 \quad (16-18)$$

Thus, the thermodynamic temperature is proportional to the average kinetic energy of the system. Moreover, at any time t of a molecular dynamics trajectory, the

instantaneous temperature T_{inst} is defined in terms of the instantaneous kinetic energy K_{inst} as

$$T_{\text{inst}} = \frac{1}{3Nk_{\text{B}}} \sum_{i=1}^N m_i v_i^2 = \frac{2}{3Nk_{\text{B}}} K_{\text{inst}} \quad (16-19)$$

The thermodynamic temperature is the average of the instantaneous temperature over many molecular dynamics time steps.

Temperature is treated differently for the various ensembles simulated. For a microcanonical (constant NVE) ensemble, the temperature fluctuates about its average value. For a canonical ensemble (constant NVT), various methods have been developed to maintain a constant temperature. In some of these methods, a heat bath is introduced with one or several extra degrees of freedom so that averages in the microcanonical ensemble for the extended system are equivalent to averages in the canonical ensemble for the real system. For an isothermal–isobaric ensemble (constant NPT), methods have been developed to maintain constant temperature and pressure simultaneously.

For the simulation of biological systems such as proteins and nucleic acids, typically the initial coordinates are obtained from an experimentally derived structure, often from X-ray crystallography or NMR. The initial velocities are usually chosen in a manner that satisfies the Maxwell–Boltzmann distribution of velocities for the specified temperature (related to Eq. (15-5)). For simulations of equilibrium properties, it is important to follow a careful equilibration procedure for the system prior to data collection. The purpose of the equilibration procedure is to relax bad contacts that may have been present in the initial structure and to allow the system to gradually settle to the desired temperature. During this equilibration procedure, the temperature can be controlled by scaling the velocities to satisfy the Maxwell–Boltzmann distribution for the desired temperature or with more sophisticated techniques. After equilibration, the physical properties of the system should fluctuate around their average values, which should remain constant over time.

Several useful quantities can be calculated from molecular dynamics simulations. The root-mean-square deviation (RMSD) quantifies the differences between the atomic coordinates for two configurations of the same protein. Often the RMSD is calculated between the configuration at a given time step and a reference configuration, such as an X-ray crystal structure. The RMSD is defined as

$$\text{RMSD} = \sqrt{\frac{1}{N} \sum_{i=1}^N |\mathbf{r}_i(t) - \mathbf{r}_i^0|^2} \quad (16-20)$$

where N is the number of atoms in the system, $\mathbf{r}_i(t)$ is the position coordinate of atom i at time t , and $\mathbf{r}_i^0$ is the position coordinate of atom i in the reference configuration. For a molecular dynamics trajectory, typically the RMSD is calculated for each time step and is averaged over all atoms of interest. Often only the heavy atoms are included in the summation; in other words, the hydrogen atoms are omitted. The RMSD for a

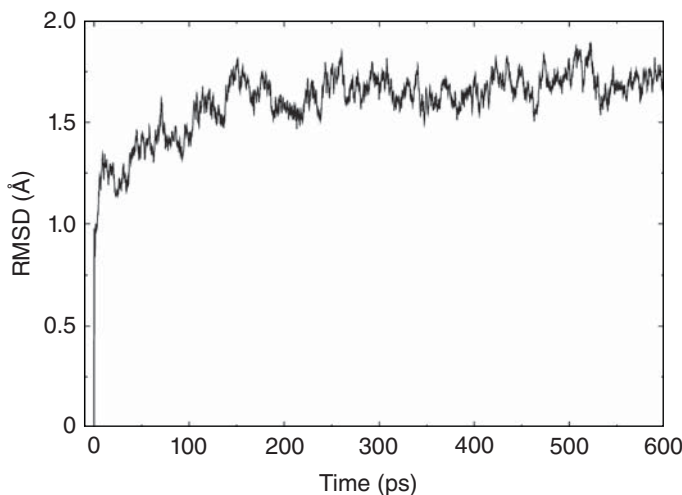


FIGURE 16-8. The RMSD, given in Eq. (16-20), for all heavy atoms in DHFR for each time step in a molecular dynamics trajectory with respect to the initial coordinates, which were obtained from an X-ray crystal structure. The RMSD changes quickly in the initial stage of the trajectory, but after a relatively short equilibration period, the RMSD fluctuates regularly about a value of 1.7 Å.

simulation of the enzyme DHFR is depicted in Figure 16-8. Here the RMSD is calculated relative to the initial coordinates, which were obtained from an X-ray crystal structure. At the beginning of the trajectory, the RMSD increases dramatically, but after a short equilibration period, the RMSD fluctuates regularly about its average value of 1.7 Å, indicating that the equilibrated structure is stable and does not deviate significantly from the X-ray crystal structure.

Another quantity that is often calculated is the root-mean-square fluctuation (RMSF), which quantifies the fluctuations of each atom in the protein about its average position during a molecular dynamics trajectory. Thus, the RMSF identifies the more flexible regions of a protein. The RMSF is defined as

$$\text{RMSF} = \sqrt{\frac{1}{T} \sum_{t=1}^T |\mathbf{r}_i(t) - \bar{\mathbf{r}}_i|^2} \quad (16-21)$$

where T is the number of time steps, $\mathbf{r}_i(t)$ is the position coordinate of atom i at time step t , and $\bar{\mathbf{r}}_i$ is the average position coordinate of atom i over the trajectory. (For notational simplicity, here t is the time step rather than the actual time.) In contrast to the RMSD, the RMSF is calculated for each atom and is averaged over all time steps. The RMSF for a simulation of DHFR is depicted in Figure 16-9. Here the RMSF is calculated only for the C_α atoms of the protein backbone. The peaks in this plot indicate the more flexible regions of DHFR. For this protein, the residues with larger RMSFs are located in either disordered loop regions or tail regions at the ends of the peptide chain.

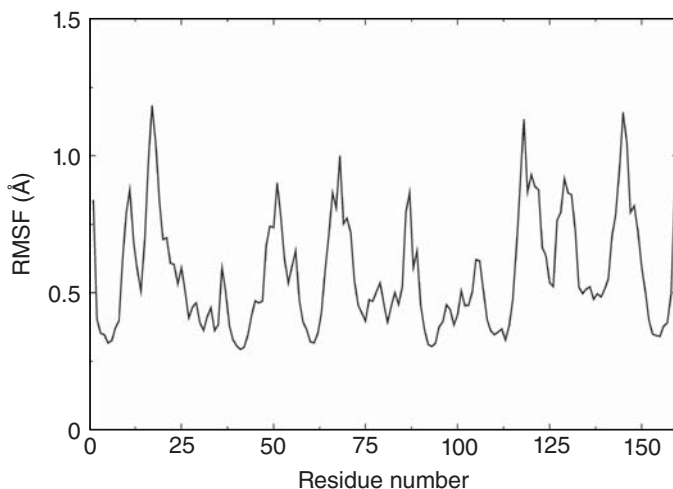


FIGURE 16-9. The RMSF, given in Eq. (16-21), for the C_{α} atoms in DHFR, averaged over a 5 ns molecular dynamics trajectory. The peaks indicate regions of greater flexibility, whereas the lower RMSF values indicate regions that are more rigid.

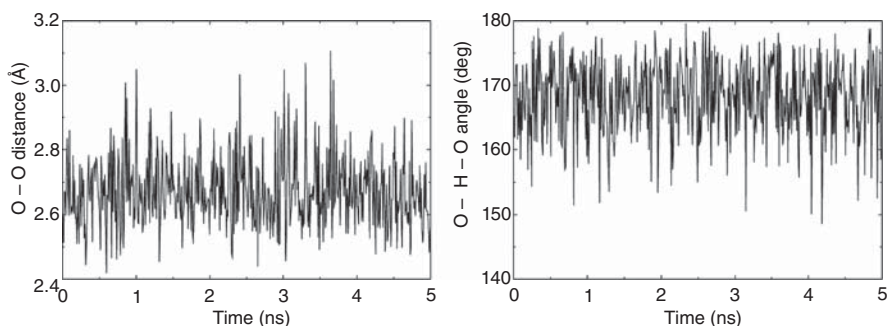


FIGURE 16-10. Geometric description of the hydrogen bond between the hydroxyl group of Thr46 and a nonbridging phosphate oxygen in the NADPH cofactor during a molecular dynamics simulation of DHFR. According to standard criteria used to define hydrogen bonds, the hydrogen bond is formed for 95% of the configurations sampled in this trajectory.

In addition to the RMSD and RMSF, structural properties of the protein can also be analyzed with molecular dynamics. For example, Figure 16-10 depicts an analysis of the hydrogen bond between the hydroxyl group of Thr46 and a nonbridging phosphate oxygen in the NADPH cofactor. The hydrogen bond donor–acceptor distance and the donor–hydrogen–acceptor angle are plotted as a function of time. The figure shows that the distance and angle fluctuate regularly about their average values. Typically a hydrogen bond is defined by criteria such as the donor–acceptor distance is no more than 3.0–3.5 Å and the donor–hydrogen–acceptor angle does not deviate

from linearity by more than $\sim 30^\circ$. Using such criteria, the hydrogen bond is formed for 95% of the configurations sampled in this trajectory. These types of analyses provide a measure of the stability, pervasiveness, and fluctuations of specific hydrogen bonds. Other types of interactions and structural properties can be examined in an analogous way.

16.6 MONTE CARLO

The Monte Carlo method was named after the famous gambling city in Europe because of the extensive use of random numbers in the algorithm. The objective of this method is to generate a set of states in phase space that are sampled from a specified statistical ensemble. In contrast to molecular dynamics, the states are not connected to each other in a physically meaningful way, and time-dependent information is not provided. Nevertheless, the Monte Carlo method is useful for the efficient calculation of average properties for a given ensemble.

The Metropolis Monte Carlo method, named after Nicholas Metropolis, a Greek American physicist, provides an algorithm for calculating average properties within a canonical ensemble (constant NVT). Starting with an initial configuration, one of the atoms is perturbed by a random displacement in three-dimensional space. The potential energy is calculated for the initial state m and for the perturbed state n , and the corresponding potential energies U_m and U_n are compared to determine whether to accept or reject this move. If the move is downhill in energy, $U_n - U_m \leq 0$, the move is accepted. If the move is uphill in energy, $U_n - U_m > 0$, the move is accepted with probability $\exp[-(U_n - U_m)/(k_B T)]$ using an algorithm that depends on a selected random number. If the uphill move is rejected, then the configuration associated with the initial state m becomes the new state in the chain. This procedure is repeated to generate a chain of states that are sampled from a canonical ensemble. The properties of the system can be averaged over this chain of states in the same way that they are averaged over molecular dynamics trajectories, although no time-dependent information is provided by the Monte Carlo method. The Monte Carlo method has also been extended to calculate average properties in other ensembles, such as the isobaric-isothermal ensemble (constant NPT).

16.7 HYBRID QUANTUM/CLASSICAL METHODS

The standard molecular mechanical force fields described in Section 16.2 do not allow chemical bonds to break or form. A variety of approaches have been developed to address this limitation. One approach is the development of reactive force fields. In this approach, a molecular mechanical force field is augmented with analytical functional terms that allow the breaking and forming of chemical bonds, and the parameters in these terms are fit to quantum chemistry calculations for a specific type of reactive system. A disadvantage of this approach is that the fitting procedure is often cumbersome and must be repeated for each different type

of system. Another approach for simulating chemical reactions is called *direct dynamics* or *ab initio dynamics*. In this approach, the entire system is treated quantum mechanically, and the potential energy surface is generated “on the fly” by performing an electronic structure calculation at each molecular dynamics time step. The main drawback of this approach is the computational expense, which severely limits the size of the system and the duration of the molecular dynamics trajectory.

Mixed quantum mechanical/molecular mechanical (QM/MM) methods provide an attractive alternative to these approaches. In QM/MM methods, the system is divided into reactive and nonreactive regions. The reactive region is treated quantum mechanically, while the nonreactive region is treated with a standard molecular mechanical force field. The QM region can be described using a variety of electronic structure methods, including semiempirical methods, Hartree–Fock theory, and density functional theory, which were discussed in Chapter 8. Various strategies have been devised to treat the boundary between the QM and MM regions. For example, the bonds that are broken at this boundary may be capped with hydrogen atoms. Moreover, numerous schemes have been developed to describe the interactions between the QM and MM regions. In general, these QM-MM interactions include both electrostatic and van der Waals interactions. More specifically, the electronic structure calculation for the QM region includes electrostatic interactions with the atomic point charges in the MM region.

Another approach that is often categorized as a QM/MM method is the empirical valence bond (EVB) method. The EVB method describes a chemical reaction in terms of a small number of resonance structures, also called valence bond states. For example, a hydrogen transfer reaction may be described in terms of the two states $AH \cdots B$ and $A \cdots HB$, where the hydrogen is covalently bonded to A in the first state and to B in the second state. These valence bond states are then used as basis functions for solving the electronic Schrödinger equation, as described in Chapter 8. To avoid explicitly solving the Schrödinger equation, the energies of these two states and the coupling between them are usually represented as molecular mechanical force field terms fit to electronic structure calculations or experimental data. Then the potential energy surface for the chemical reaction is expressed as a linear combination of these force field terms. Thus, the EVB method is often nearly as computationally efficient as standard molecular mechanical force fields but still allows chemical bonds to break and form. The disadvantage of this approach is the required parameterization procedure, although methods have been developed to automate or avoid this procedure by connecting the valence bond approach directly to electronic structure methods.

In addition to electronic quantum mechanical effects, which are required to describe bond breaking and forming, the nuclei can also exhibit quantum mechanical effects. Nuclear quantum effects that can be important in biological systems include zero-point energy and tunneling, as discussed in Chapter 7. In particular, hydrogen tunneling has been shown to play an important role in a variety of proton and hydride transfer reactions catalyzed by enzymes. Various methods have been developed to

include nuclear quantum effects in molecular simulations, but a discussion of these methods is beyond the scope of this book.

16.8 HELMHOLTZ AND GIBBS ENERGY CALCULATIONS

Molecular dynamics simulations can also be used to calculate relative Helmholtz and Gibbs energies. Helmholtz energies are obtained from the canonical ensemble (constant NVT), while Gibbs energies are obtained from the isothermal–isobaric ensemble (constant NPT). Most biological simulations are performed in the solution phase, where these two thermodynamic quantities are very similar and often are not even distinguished. In the literature, these quantities are usually referred to as *free energies*, and this term will be used in this section to denote either Helmholtz or Gibbs energies.

Using the statistical mechanical concepts discussed in Chapter 15, the relative free energies between two states, denoted A and B, can be calculated from the partition functions for the two states. When these two states are significantly different, technical difficulties arise because large energy differences between the two states lead to small values of the exponential in the partition function, thereby causing statistical uncertainty and numerical errors. To address this challenge, a variety of simulation methods have been developed to calculate relative free energies. These methods involve constructing a path connecting states A and B in conformational space and breaking this path up into windows, where each window is small enough to allow the accurate calculation of the free energy change across that window. Because the free energy is a state function, the free energy difference between the end points does not depend on the particular path chosen, and often this path has no physical meaning. Various methods have been developed to bias the potential energy surface to ensure adequate sampling within each window. The data from all of the windows are combined using statistical methods that provide the free energy difference between the end points. These types of approaches can be used to obtain information about the free energy differences associated with a conformational change, ligand binding, and other biological transformations.

In addition, related methods have been developed for calculating the free energy barriers for chemical reactions in biological systems. Many of these methods are based on *umbrella sampling*, in which an additional term is added to the potential energy surface to bias the system toward a particular region of conformational space. In practice, a series of independent molecular dynamics trajectories is propagated with different biasing potentials that cause the system to evolve along a proposed reaction coordinate. Subsequently, all of the data are combined using statistical methods that remove the effects of the biasing potentials to obtain the free energy along this reaction coordinate for the original potential energy surface. The choice of reaction coordinate can significantly influence the calculated free energy barrier. For this reason, methods have been developed to generate a multidimensional free energy surface along several reaction coordinates that are chosen to represent the key interatomic distances or angles. Moreover, in certain cases, a collective reaction coordinate

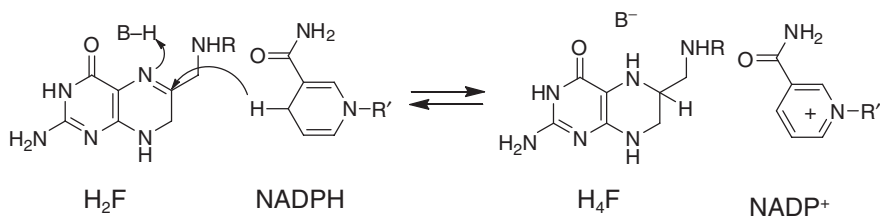


FIGURE 16-11. The proton and hydride transfer reactions catalyzed by the enzyme DHFR. BH is a proton donor, e.g., the enzyme or H_2O .

that depends on a large number of nuclear coordinates is more suitable than a single geometrical reaction coordinate. An example of a free energy calculation along this type of collective reaction coordinate will be given in the next section for the hydride transfer reaction catalyzed by the enzyme DHFR.

16.9 SIMULATIONS OF ENZYME REACTIONS

As discussed earlier in this book (Chapters 10 and 12), the enzyme dihydrofolate reductase (DHFR) catalyzes the conversion of 7,8-dihydrofolate (DHF) to 5,6,7,8-tetrahydrofolate (THF), using nicotinamide adenine dinucleotide phosphate (NADPH) as a coenzyme. This mechanism is thought to proceed through the initial protonation of DHF, followed by hydride transfer from the NADPH cofactor to the protonated DHF substrate. The hydride transfer reaction catalyzed by DHFR is depicted in Figure 16-11. This reaction has been studied extensively with an array of experimental methods, including X-ray crystallography, NMR, and single molecule spectroscopy.

This enzymatic reaction has also been studied with a wide range of theoretical approaches. The potential energy surface has been represented by several different QM/MM and EVB methods that have provided predominantly consistent results. The approaches described in the previous section have been used to calculate the free energy barrier for the hydride transfer reaction and to investigate the conformational changes that occur along the reaction pathway. In particular, the distances between key residues have been monitored during this reaction. Moreover, the impact of mutating specific amino acids on the free energy barrier and on the conformational changes along the reaction pathway has also been studied with these approaches. The corresponding changes in rate constants calculated within the framework of transition state theory are in agreement with experimental results. In addition, several different approaches for including the nuclear quantum effects of the transferring hydrogen have been applied to this enzyme, and the calculated kinetic isotope effects associated with replacing the transferring hydrogen with deuterium are in agreement with experimental values. This agreement between theoretical calculations and experimental data provides validation for the computational methods.

A combination of theoretical and experimental data has provided insights into the role of protein motion and conformational sampling in enzyme catalysis. The

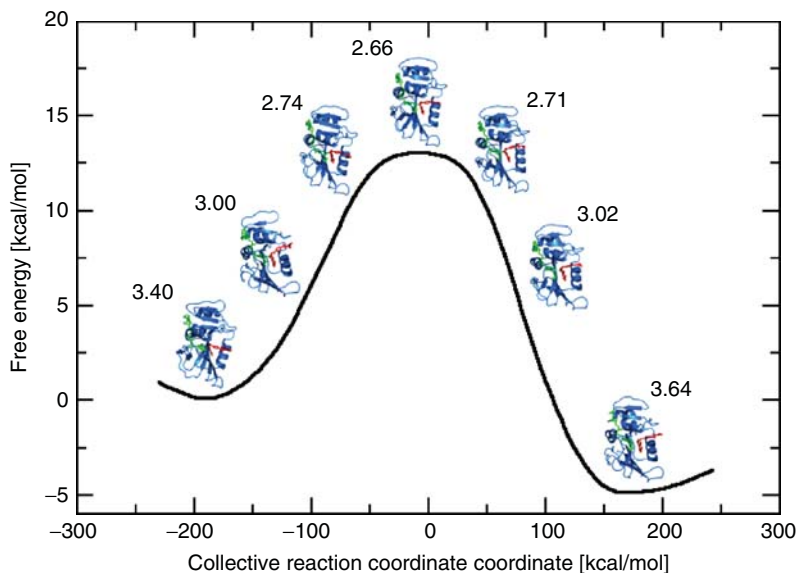


FIGURE 16-12. Free energy profile of the DHFR-catalyzed hydride transfer reaction as a function of a collective reaction coordinate defined in terms of the energy difference between the reactant and product valence bond states. This reaction coordinate includes motions of the enzyme, substrate, and cofactor. The magnitude of the free energy barrier is determined by the relative probabilities of sampling configurations at the top of the free energy barrier and in the reactant minimum. The average donor–acceptor distances in Angstroms are provided next to the free energy curve for selected values of the reaction coordinate. Note that the donor–acceptor distance decreases as the reaction evolves from the reactant to the top of the barrier. The conformational changes along the collective reaction coordinate are attained by equilibrium, stochastic, thermal motions occurring within the confines of the protein structure. These conformational changes facilitate the hydride transfer reaction by bringing the donor and acceptor closer together, orienting the substrate and cofactor properly, and providing a favorable electrostatic environment. Reproduced in part with permission from G. G. Hammes, S. J. Benkovic, and S. Hammes-Schiffer, *Biochemistry* **50**, 10422 (2011). © 2011 by American Chemical Society.

concepts are summarized in Figure 16-12 and Figure 16-13. According to this perspective, stochastic thermal motions throughout the enzyme and ligands lead to conformational sampling of configurations that facilitate hydride transfer by bringing the substrate and cofactor closer together, orienting them properly, and providing a favorable electrostatic environment. These thermal motions are Brownian in nature but occur within the confines of the protein structure. Typically these thermal motions occur on the pico- to nanosecond time scale but lead to overall conformational changes that occur on a slower timescale along the collective reaction coordinate. Within this framework, the experimentally measured millisecond timescale for hydride transfer in DHFR is related to the average time required to sample configurations that are conducive to the chemical reaction. The actual chemical bond

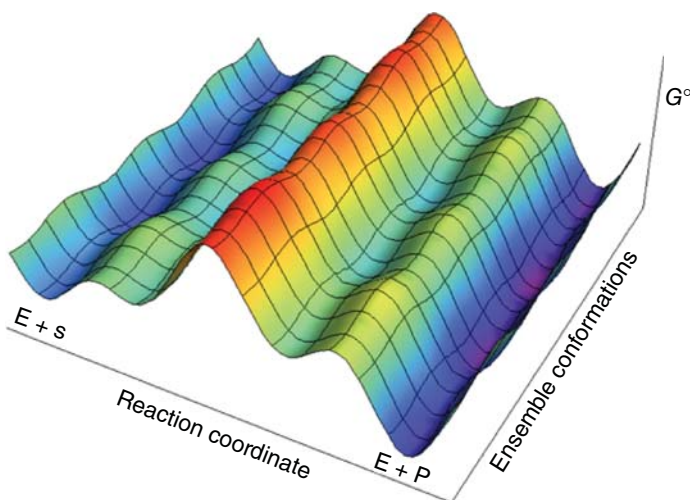


FIGURE 16-13. Schematic representation of the standard free energy landscape for the catalytic network of an enzyme reaction. Conformational changes occur along both axes. The conformational changes occurring along the Reaction Coordinate axis correspond to the environmental reorganization that facilitates the chemical reaction. In contrast, the conformational changes occurring along the Ensemble Conformations axis represent the ensembles of configurations existing at all stages along the reaction coordinate, leading to a large number of parallel catalytic pathways. The reaction paths can slide along and between both coordinates. For real enzymes, the number of maxima and minima along the coordinates is expected to be greater than shown. The free energy landscape and dominant catalytic pathways will be altered by external conditions and protein mutations. Reproduced in part with permission from G. G. Hammes, S. J. Benkovic, and S. Hammes-Schiffer, *Biochemistry* **50**, 10422 (2011). © 2011 by American Chemical Society.

breaking and forming occurs much faster than this millisecond timescale and can be viewed as a rare event that occurs only after sufficient conformational sampling generates a configuration conducive to the chemical reaction.

The impact of mutations on the hydride transfer rate constant for DHFR can be understood in terms of the free energy barrier along the collective reaction coordinate for hydride transfer. As discussed in Chapter 15, the free energy barrier, or standard Gibbs energy barrier, can be viewed as the relative probability of sampling configurations at the top of the barrier compared to those in the reactant minimum. Mutations alter the potential energy surface on which the protein and ligands move, thereby altering the relative probabilities of sampling these regions of the multidimensional surface. By changing the equilibrium conformational sampling of the system, the mutation changes the free energy barrier and therefore the rate constant. This perspective has significant implications for protein engineering and drug design. In addition to structural changes, the impact of mutations on the equilibrium conformational motions should also be considered.

Thus, molecular simulations of DHFR have provided insights into the fundamental principles underlying enzyme catalysis. A wide range of other enzymes have also been studied with these types of computational methods. As the algorithms are improved and computers become faster, further insights will be obtained from molecular simulations. These types of advances will enable the computational design of highly efficient and specific enzymes and inhibitors for both medical and technological purposes.

REFERENCES

The topics in this chapter are discussed in considerably more depth in a number of textbooks, such as Refs. 1 and 2. A more detailed discussion of issues pertaining to simulations of enzyme reactions is provided in Refs. 3–6.

1. M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids*, Oxford University Press, New York, NY, 1987.
2. A. Warshel, *Computer Modeling of Chemical Reactions in Enzymes and Solutions*, John Wiley & Sons, Inc., New York, NY, 1991.
3. S. Hammes-Schiffer and S. J. Benkovic, *Ann. Rev. Biochem.* 75, 519–541 (2006).
4. S. J. Benkovic, G. G. Hammes, and S. Hammes-Schiffer, *Biochemistry* 47, 3317–3321 (2008).
5. G. G. Hammes, S. J. Benkovic, and S. Hammes-Schiffer, *Biochemistry* 50, 10422 (2011).
6. S. Hammes-Schiffer, *Biochemistry* 52, 2012–2020 (2013).

PROBLEMS

- 16-1.** Consider two atoms, labeled A and B, interacting according to the van der Waals interaction given in Eq. (16-10). The minimum interaction energy occurs at a separation of $R_{\min}$. Assume there are no other atoms in the system.
- a. The two atoms are separated by $R_{\min} + 0.5 \text{ \AA}$. Is the interaction between the atoms attractive or repulsive at this distance? Does the force on the atoms pull them toward each other or push them away from each other?
 - b. The two atoms are separated by $R_{\min} - 0.5 \text{ \AA}$. Is the interaction between the atoms attractive or repulsive at this distance? Does the force on the atoms pull them toward each other or push them away from each other?
 - c. The two atoms are separated by $R_{\min}$. Is the interaction between the atoms attractive or repulsive at this distance? Does the force on the atoms pull them toward each other or push them away from each other?
- 16-2.** Consider a molecular dynamics simulation of liquid ethane (C_2H_6). The following questions relate to the molecular mechanical force field.

- a. How many bonds, angles, and dihedrals (torsion angles) should be defined to characterize the topology of a single ethane molecule? Here you need to count all of the possible bonds, three-atom angles, and torsion angles. All of the corresponding bonding interaction energies for each ethane molecule are included in the force field.
 - b. There are only two atom types in the ethane molecule (“C” and “H”), and each atom type is assigned van der Waals parameters and a partial charge to describe the nonbonding interactions. The nonbonding interaction energies for all pairs of ethane molecules are included in the force field. Consider the intermolecular interaction between two ethane molecules. How many nonbonding interactions between two carbon atoms should be included? How many nonbonding interactions between two hydrogen atoms should be included? How many nonbonding interactions between a carbon and a hydrogen atom should be included? In each case, a single nonbonding interaction is assumed to include both van der Waals and Coulomb interactions, and only the *intermolecular* nonbonding interactions between two ethane molecules should be counted.
 - c. Consider the dihedral angle H1-C1-C2-H2 in an ethane molecule, where H1 is bonded to C1 and H2 is bonded to C2. What is the periodicity of the corresponding dihedral angle potential U_ϕ ? Sketch the plot of the potential for this dihedral as a function of the angle changing from 0 to 360 degrees. Here each minimum of the potential should correspond to the energetically most favorable conformation of the molecule.
- 16-3.** Consider a single particle of a unit mass $m = 1$ moving in a one-dimensional harmonic potential $U(x) = \frac{1}{2}(x - 1)^2$. At time $t = 0$, the position of the particle is $x(0) = 0$ and the velocity of the particle is $v(0) = 0$. Assume that all units are consistent and that you do not need to carry them through the problem.
- a. Calculate the initial force $F(0)$ acting on the particle at $t = 0$.
 - b. Propagate the position and velocity of the particle for a single time step $\Delta t = 0.5$ using the velocity Verlet algorithm given in Eqs. (16-14) and (16-15). The acceleration can be calculated with Eq. (16-12) with the force calculated from Eq. (16-13).
 - c. The total energy of the particle is the sum of the kinetic and potential energies. Calculate the total energy of the particle at initial time $t = 0$ and at time $t = 0.5$ after the first propagation step.
 - d. In an isolated system, the total energy should be conserved along a trajectory. Based on your answer from the previous part, was the total energy conserved during the single time step you propagated? If not, explain why not.
- 16-4.** Consider the equilibration stage of a molecular dynamics simulation of a protein in water. At a certain time step, the instantaneous temperature is T_{inst} . If the

target temperature is T_{target} , by what factor should you scale all of the velocities so that the new instantaneous temperature is T_{target} ? This factor should be independent of mass.

- 16-5.** The hydride transfer reaction catalyzed by DHFR was simulated using an EVB potential energy surface with the umbrella sampling method described earlier to generate the free energy curve along a collective reaction coordinate at 298 K. When the residue Gly121 was mutated to Val, the standard Gibbs energy barrier increased by 3.4 kcal/mol. Using transition state theory, estimate the change in rate constant that this simulation predicts will be observed for this mutant relative to the wild-type enzyme.

SPECIAL TOPICS

Ligand Binding to Macromolecules

17.1 INTRODUCTION

The binding of ligands to macromolecules is a key element in virtually all biological processes. *Ligands* can be small molecules, such as metabolites, or large molecules, such as proteins and nucleic acids. Ligands bind to a variety of *receptors*, such as enzymes, antibodies, DNA, and membrane-bound proteins. For example, the binding of substrates to enzymes initiates the catalytic reaction. The binding of hormones, such as insulin, to receptors regulates metabolic events, and the binding of repressors and activators to DNA regulates gene transcription. The uptake and release of oxygen by hemoglobin is essential for life. Indeed, compilation of a comprehensive list of biological processes in which ligand binding plays a key role would be a formidable task.

In this chapter, we shall discuss how to analyze ligand binding to macromolecules quantitatively for both simple and complex systems. We will also consider experimental methods that are used to study ligand binding. The application and importance of this analysis for biology will be illustrated through specific examples. The treatment presented will be adequate for most situations: more complete (and more complex) discussions of this topic are available (1–3).

17.2 BINDING OF SMALL MOLECULES TO MULTIPLE IDENTICAL BINDING SITES

The binding of a small molecule to identical sites on a macromolecule is a common occurrence. For example, enzymes frequently have several binding sites for substrates on a single molecule. Proteins have multiple binding sites for protons that are often essentially identical, for example, carboxyl or amino groups. Let us assume the simplest case, namely, a single ligand, L, binding to a single site on a protein, P:



This equilibrium can be characterized by the equilibrium constant, K :

$$K = \frac{[PL]}{[P][L]} \quad (17-2)$$

Binding equilibria involving macromolecules are conveniently characterized by a binding isotherm, the moles of ligand bound per mole of protein, r . For the aforementioned case,

$$r = \frac{[PL]}{[P] + [PL]} = \frac{K[L]}{1 + K[L]} \quad (17-3)$$

A plot of r versus $[L]$ is shown in Figure 17-1a: it is a hyperbolic curve with a limiting value of 1 at high ligand concentrations, and when $r = 0.5$, $[L] = 1/K$.

If there are n identical binding sites on the protein, the binding isotherm for the macromolecule is simply the sum of those for each of the sites:

$$r = \frac{nK[L]}{1 + K[L]} \quad (17-4)$$

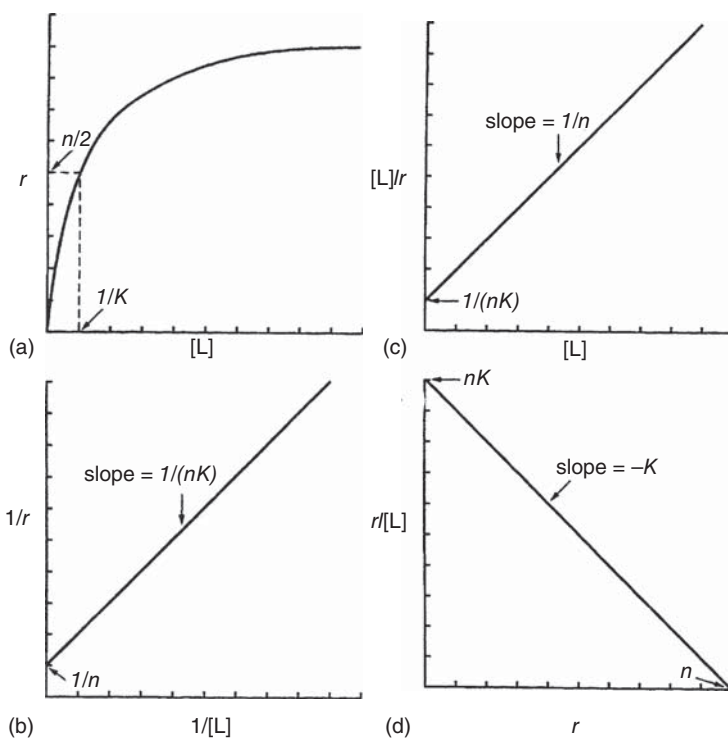


FIGURE 17-1. Plots of binding isotherms for n identical sites on a macromolecule according to (a) Eq. (17-4), (b) Eq. (17-5), (c) Eq. (17-6), and (d) Eq. (17-7). When $n = 1$, the curve in (a) obeys Eq. (17-3).

A plot of r versus $[L]$ has the same shape as before, but now the limiting value of r at high ligand concentrations is n , the number of identical binding sites on the macromolecule, and when $r = n/2$, $[L] = 1/K$. Thus, a study of ligand binding to a protein containing n identical binding sites permits determination of the number of binding sites and the equilibrium binding constant. In practice, alternative plots of the data are frequently used that yield a straight line plot and therefore can be analyzed more easily. With the availability of nonlinear least squares programs on desktop computers, this is not really necessary. However, it should be kept in mind that statistical analyses always fit the data—this does not necessarily mean that the fit is a good one. The quality of the fit must carefully be examined to be sure that the equations used and the fitting procedures are appropriate. In addition, it is extremely important that a very wide range of ligand concentration is used.

The most obvious recasting of Eq. (17-4) is to take its reciprocal:

$$\frac{1}{r} = \frac{1}{n} + \frac{1}{nK[L]} \quad (17-5)$$

A plot of $1/r$ versus $1/[L]$ is a straight line with an intercept on the $[L]$ axis of $1/n$ and a slope of $1/(nK)$. Another possibility is to multiply Eq. (17-5) by $[L]$:

$$\frac{[L]}{r} = \frac{[L]}{n} + \frac{1}{nK} \quad (17-6)$$

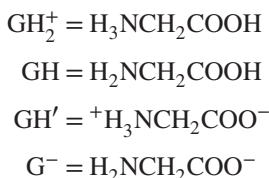
If $[L]/r$ is plotted versus $[L]$, a straight line is obtained with a slope of $1/n$ and an intercept of $1/(nK)$. A common alternative plot is a Scatchard plot (4), named after George Scatchard, a pioneer in the study of small molecule binding to proteins. Rearrangement of Eq. (17-4) gives

$$\frac{r}{[L]} = nK - rK \quad (17-7)$$

A plot of $r/[L]$ versus r is a straight line, with the intercept on the x -axis ($r/[L] = 0$) being equal to the number of binding sites, n , and the intercept on the y -axis being equal to nK . Examples of these straight line plots are included in Figure 17-1.

17.3 MACROSCOPIC AND MICROSCOPIC EQUILIBRIUM CONSTANTS

Before we proceed further with the analysis of ligand binding to macromolecules, it is important to understand the distinction between macroscopic and microscopic equilibrium constants. Consider, for example, a dibasic acid such as the amino acid glycine. Four protonation states are possible:



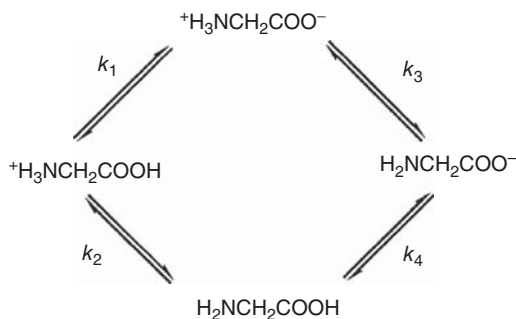
If a pH titration is carried out, the states GH and GH' cannot be distinguished as they contain the same number of protons. The pH titration can be used to determine the *macroscopic* ionization constants:

$$K_1 = \frac{([GH] + [GH'])[H^+]}{[GH_2^+]} \quad (17-8)$$

$$K_2 = \frac{[G^-][H^+]}{([GH] + [GH'])} \quad (17-9)$$

The two pK values determined by pH titration are $pK_1 = 2.35$ and $pK_2 = 9.78$ (25°C and zero salt concentration).

If we consider the *microscopic* states of glycine, four *microscopic ionization constants*, k_i , are needed to characterize the system:



Note that these four microscopic ionization constants are not independent because this is a closed cycle. If the four microscopic ionization constants are written in terms of the concentrations, it can be seen that

$$k_1 k_3 = k_2 k_4 \quad (17-10)$$

This is an example of the principle of detailed balance. Whenever a closed cycle of reactions occurs, a relationship between the individual equilibrium constants such as Eq. (17-10) exists. The relationship between the macroscopic and microscopic ionization constants can easily be seen by reference to Eqs. 17-8 and 17-9, namely,

$$K_1 = k_1 + k_2 \quad (17-11)$$

and

$$K_2 = \frac{k_3 k_4}{(k_3 + k_4)} \quad (17-12)$$

Since the two macroscopic ionization constants and Eq. (17-10) provide only three relationships between the microscopic ionization constants, it is apparent that the

microscopic ionization constants cannot be determined from a pH titration. In order to calculate the microscopic ionization constants, additional information is needed. In this case, it might be assumed that k_2 is equal to the ionization constant determined by pH titration of the methyl ester of glycine, namely, $pK_2 = 7.70$. This assumes that the ionization constant for protonation of the amino group is the same when either a proton or a methyl group is bound to the carboxyl, which is not unreasonable. With this assumption, the other three microscopic pK values can be calculated: $pK_1 = 2.35$, $pK_3 = 9.78$, and $pK_4 = 4.43$. These results indicate that the bottom state in the ionization scheme is present at very low concentrations throughout a pH titration.

This is a simple illustration of the distinction between macroscopic and microscopic equilibrium constants. Experimental measurements usually only give information about macroscopic binding constants although exceptions exist. For example, the state of protonation of the nitrogen can be monitored by nuclear magnetic resonance (NMR) so that, in principle, a titration curve can be obtained for the amino group independently of that for the dibasic acid.

17.4 STATISTICAL EFFECTS IN LIGAND BINDING TO MACROMOLECULES

Let us return again to the binding of a ligand to multiple identical binding sites on a macromolecule and consider the matter of macroscopic and microscopic equilibrium constants. As a simple example, consider a macromolecule, P, with two identical binding sites; symbolically, one of the binding sites will be on the left of P and the other on the right of P. The macroscopic equilibrium constants are

$$K_1 = \frac{([PL] + [LP])}{[L][P]} \quad (17-13)$$

and

$$K_2 = \frac{[PL_2]}{[L]([PL] + [LP])} \quad (17-14)$$

In relating this discussion to the previous consideration of glycine, it should be noted that the glycine proteolytic equilibria are characterized by equilibrium *dissociation* constants whereas equilibrium *association* constants are used here.

The equilibrium binding isotherm can be written as

$$r = \frac{[PL] + [LP] + 2[PL_2]}{[P] + [PL] + [LP] + [PL_2]} \quad (17-15)$$

If both the numerator and the denominator are divided by $[P]$, it can easily be seen that

$$r = \frac{K_1[L] + 2K_1K_2[L]^2}{1 + K_1[L] + K_1K_2[L]^2} \quad (17-16)$$

This does not look the same as Eq. (17-4). This is because the equilibrium constant in Eq. (17-4) is the microscopic equilibrium constant. In this case, it is easy to see that $K_1 = 2K$ and $K_2 = K/2$. If these relationships are put into Eq. (17-16),

$$r = \frac{2K[L](1 + K[L])}{(1 + K[L])^2}$$

$$r = \frac{2K[L]}{1 + K[L]}$$

Thus, in the case of multiple identical binding sites, the binding isotherm is the same whether macroscopic or microscopic equilibrium constants are considered. The relationship between microscopic and macroscopic constants for this particular case is particularly simple. When both sites are empty, there are two possible sites available for the ligand so that the macroscopic equilibrium constant is twice as large as the microscopic constant. When both sites are occupied by a ligand, there are two sites from which the ligand can dissociate; therefore, the macroscopic equilibrium constant is one-half of the microscopic constant.

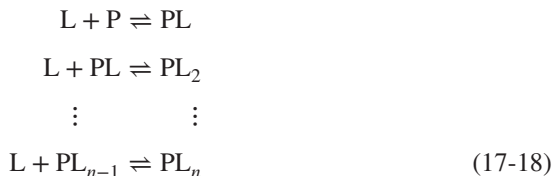
If more than two identical binding sites are present, the relationship between the macroscopic and microscopic equilibrium constants can be determined in a similar manner. If a macromolecule has n identical binding sites, the relationship between the macroscopic equilibrium constant for binding to the i th site and the microscopic equilibrium constant is

$$K_i = \frac{\text{number of free sites on P before binding}}{\text{number of occupied sites on P after binding}} K$$

$$K_i = [(n - i + 1)/i]K, \quad i \geq 1 \quad (17-17)$$

We now return to consideration of a macromolecule with n identical binding sites.

The macroscopic equilibria are



The corresponding macroscopic equilibrium constants are

$$\begin{aligned} K_1 &= [PL]/([P][L]) \\ &\vdots \\ K_n &= [PL_n]/([L][PL_{n-1}]) \end{aligned} \quad (17-19)$$

The binding isotherm can be written as

$$r = \frac{[\text{PL}] + 2[\text{PL}_2] + \cdots + n[\text{PL}_n]}{[\text{P}] + [\text{PL}] + [\text{PL}_2] + \cdots + [\text{PL}_n]} \quad (17-20)$$

If we divide the numerator and denominator by $[\text{P}]$ and use the definitions of the macroscopic association constants for the ratios $[\text{PL}_i]/[\text{P}]$, Eq. (17-20) becomes

$$r = \frac{K_1[\text{L}] + 2K_1K_2[\text{L}]^2 + \cdots + nK_1K_2 \cdots K_n[\text{L}]^n}{1 + K_1[\text{L}] + K_1K_2[\text{L}]^2 + \cdots + K_1K_2 \cdots K_n[\text{L}]^n} \quad (17-21)$$

This equation describes the binding of a ligand to n *different* binding sites as no relationship has been assumed between the binding constants. If the sites are identical, the macroscopic constants are related to the microscopic or intrinsic binding constant by Eq. (17-17). If this relationship is inserted into Eq. (17-21), it becomes

$$\begin{aligned} r &= \frac{nK[\text{L}] + \frac{2n(n-1)}{2!}K^2[\text{L}]^2 + \cdots + nK_n[\text{L}]^n}{1 + nK[\text{L}] + \frac{n(n-1)}{2!}K^2[\text{L}]^2 + \cdots + K_n[\text{L}]^n} \\ r &= \frac{nK[\text{L}](1 + K[\text{L}])^{n-1}}{(1 + K[\text{L}])^n} = \frac{nK[\text{L}]}{1 + K[\text{L}]} \end{aligned} \quad (17-22)$$

The binomial theorem has been used to obtain the final result; that is, the series of terms in the denominator that contain successive powers of $[\text{L}]$ is recognized as the expansion of $(1 + K[\text{L}])^n$. Similarly, the expansion of $(1 + K[\text{L}])^{n-1}$ can be factored out of the numerator. This result is exactly the same as Eq. (17-4), which was obtained in a more intuitive manner.

To conclude this section, we consider the situation where a macromolecule has more than one set of independent identical binding sites for a ligand. In this case, the equilibrium binding isotherm is simply a sum of the terms given in Eq. (17-4), with the number of terms in the sum being equal to the number of sets of binding sites:

$$r = \sum \frac{n_i K_i [\text{L}]}{1 + K_i [\text{L}]} \quad (17-23)$$

Here n_i and K_i are the number of sites in each set and the intrinsic equilibrium constant for each set, respectively. Unless independent information is available about the structure of the macromolecule, it is usually preferable to use Eq. (17-21) to fit the data as no assumptions are made about the nature or number of the binding sites in this case.

Virtually any ligand binding isotherm can be fit to Eq. (17-21), but this is not a meaningful exercise unless the result can be related to the structure and/or the function of the macromolecule. For example, the binding isotherm for the binding of laurate ion to human serum albumin is shown in Figure 17-2 (5,6). The concentration range covered is so large that the concentration axis is logarithmic. The corresponding Scatchard plot is also shown in Figure 17-2. It is clear that saturation of the binding

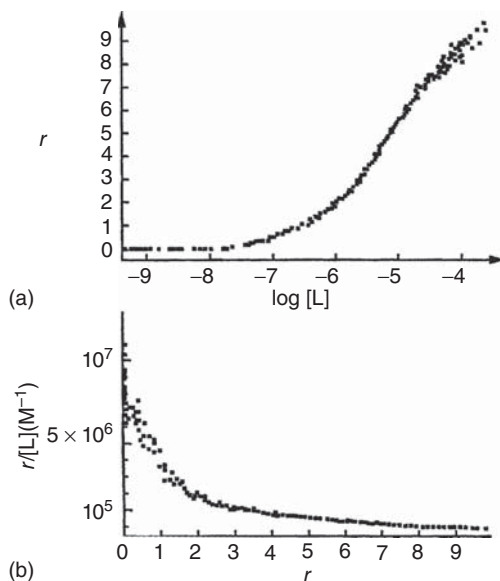


FIGURE 17-2. (a) Plot of r versus $\log [L]$ for the binding of laurate ion by human serum albumin. (b) The same data plotted as $r/[L]$ versus r (Scatchard plot). Adapted from I.M. Klotz, *Ligand-Receptor Complexes*, John Wiley & Sons, Inc. New York, 1997. Copyright © 1997 John Wiley & Sons, Inc. Reprinted by permission of John Wiley & Sons, Inc.

sites on serum albumin is never reached. The data in the figures, nevertheless, can be fit well with the assumption that $n = 10$. Clearly, this does not establish that 10 binding sites exist, nor regrettably does it provide information about the structure of the macromolecule. We will not dwell on this matter here. We will, however, return to the topic of multiple binding sites when discussing the concept of cooperativity.

17.5 EXPERIMENTAL DETERMINATION OF LIGAND BINDING ISOTHERMS

We will now make a digression to discuss briefly some of the experimental aspects of ligand binding studies. Many different experimental methods exist for determining the binding of ligands to macromolecules. The experimental methods used fall into two classes: (i) direct determination of the unbound ligand concentration and (ii) measurement of a change in physical or biological property of the ligand or macromolecule when the ligand binds. Generally, stoichiometry and binding constants can be determined reliably only if concentrations of the unbound ligand, the bound ligand, and the unbound macromolecule can be determined.

The most straightforward method is equilibrium dialysis. This method is pictured schematically in Figure 17-3. With this method, a solution is separated into two parts by a semipermeable membrane, which will not permit the macromolecule to cross

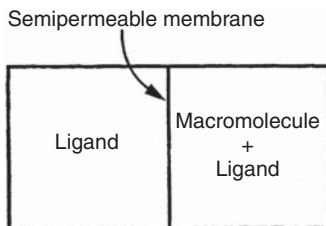


FIGURE 17-3. Schematic representation of an equilibrium dialysis experiment. The macromolecule is on one side of a semipermeable membrane and cannot pass through the membrane. The ligand can pass through the membrane and therefore is on both sides of the membrane. The concentration of the ligand on the side of the membrane that does not contain the macromolecule is equal to the concentration of the unbound ligand.

but will permit the ligand to pass freely. The macromolecule is put on one side of the membrane and the ligand on both sides. The system is then permitted to come to equilibrium. At equilibrium, the concentration of the unbound ligand is the same on both sides of the membrane. (Strictly speaking, their thermodynamic activities are equal, but we will not worry about the difference between activity and concentration here.) If the total amount of ligand is known, and the unbound amount of ligand is known, the amount of bound ligand can be determined from mass balance. The concentration of the macromolecule can be determined independently so that the binding isotherm can be calculated directly.

If equilibrium dialysis is to provide reliable data for analysis of the binding isotherm, it is usually necessary for the concentrations of bound and unbound ligand and macromolecule to be comparable. If essentially all of the ligand is bound, the amounts of unbound ligand and/or protein become very small and difficult to determine. On the other hand, if essentially no ligand is bound, calculation of the bound ligand becomes problematic. In the former case, the stoichiometry of binding can often be determined, but not the binding constant(s). In the latter case, it is virtually impossible to determine either the stoichiometry or the binding constant(s). Another way of stating this is that the concentrations of all the species should be the same order of magnitude as the reciprocal of the binding constant so that experimental points can be determined over a wide range of r . This limitation becomes restrictive for very weak binding, primarily in terms of the large amount of material required. In the case of very tight binding, it may be difficult to work with the very dilute solutions required.

Equilibrium dialysis is an example of the more general method of determining the distribution between phases. This could, for example, be the equilibration between two nonmiscible solvents that partition the reactants, although this is rarely used. Various gel exclusion media such as Agarose are, however, often used. Such gels will exclude macromolecules from their beads, with the size of the macromolecule excluded depending on the property of the specific gel. The gel will include the ligand so that the gel bead serves as a semipermeable membrane, with equilibration occurring between the inside and the outside of the bead. If a macromolecule is passed through a column equilibrated with a given ligand concentration, the unbound ligand

concentration is equal to this concentration. The amount of ligand bound and the macromolecule concentration can be determined by analysis of the column effluent. Many variations on this type of experimental protocol exist.

In some cases, the concentration of the unbound ligand is determined directly in the absence of a semipermeable membrane, for example, by a biological assay. In this case, the implicit assumption is that the rate of equilibration of the ligand and macromolecule occurs more slowly than the time for analysis. This is not necessarily the case, so such methods must be used with extreme caution.

Perturbations in the physical properties of the ligand are often used to determine binding isotherms. However, some care and caution must be taken in doing so. The most frequent situation is when the optical absorption properties of the ligand are different for the bound and unbound ligand, or for the macromolecule with and without ligand. The technique of difference spectroscopy is often used (cf. Chapter 10). With this method, the difference in the optical absorption of a solution containing the ligand and a solution containing the same concentration of ligand plus the macromolecule is determined. If the concentration of the total ligand in the first solution is $[L_T]$ and the concentrations of the free (unbound) ligand and the bound ligand are $[L_f]$ and $[L_b]$, respectively, this difference, ΔA , can be written as

$$\begin{aligned}\Delta A &= \epsilon_f[L_T] - \epsilon_f[L_f] - \epsilon_b[L_b] \\ \Delta A &= (\epsilon_f - \epsilon_b)[L_b]\end{aligned}\tag{17-24}$$

where ϵ_i are extinction coefficients and the relationship $[L_T] = [L_f] + [L_b]$ has been used. This derivation assumes that the extinction coefficient of bound ligand is the same at all binding sites, which may or may not be valid. Consequently, this method is not as direct as the partitioning methods previously described. The difference extinction coefficient, $\epsilon_f - \epsilon_b$, can be determined by extrapolation to high ligand concentrations where the macromolecule is saturated with the ligand. The difference absorbance measurements will then permit the determination of the bound ligand concentration, and the concentrations of the other species can be determined by mass balance. This method also works if the spectral change occurs in the macromolecule rather than in the ligand. The only difference is that the comparison solution contains macromolecule rather than ligand, and the concentrations and extinction coefficients of the protein appear in Eq. (17-24). Perhaps not so obvious is the fact that this method also works if spectral changes occur in both the ligand and the macromolecule when the ligand binds. The derivation of ΔA for this case is given in Chapter 10.

As a specific example of a difference spectrum titration, the difference absorbance is shown as a function of the ligand concentration for the binding of 2'-cytidine monophosphate to ribonuclease A in Figure 17-4 (7). Other optical properties such as fluorescence and circular dichroism can be used in a similar manner. Finally, it is sometimes possible to monitor a specific site on the macromolecule with physical methods such as NMR or other types of spectroscopy. If changes occur when the ligand binds, the concentrations of empty and occupied sites can be determined and translated into binding isotherms.

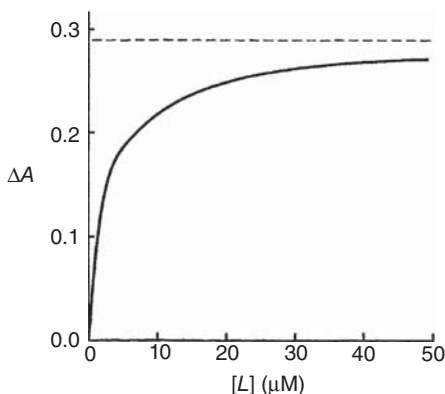


FIGURE 17-4. Plot of the absorbance difference at 288 nm, ΔA , accompanying the binding of 2'-CMP to ribonuclease A versus the concentration of unbound 2'-CMP at pH 5.5, 25°C. The total enzyme concentration is 10^{-4} M. The difference extinction coefficient is $2.88 \times 10^3 \text{ cm}^{-1} \text{ M}^{-1}$ and the binding constant is $2.96 \times 10^5 \text{ M}^{-1}$. The curve was calculated from the data in D.G. Anderson, G.G. Hammes, and Frederick G. Walz, Jr., *Biochemistry* 7, 1637 (1968). The dashed line is the difference absorbance when all of the enzyme has been converted to enzyme-2'-CMP.

The determination of high-quality binding isotherms is not trivial, and this discussion does not do full justice to the topic. One of the difficulties not discussed that should be mentioned is the occurrence of nonspecific binding. It is commonplace for nonspecific binding to occur along with binding of the ligand to the site(s) of interest. Macromolecules have many different side chains that can attract ligands, for example, by charge-charge interactions. This nonspecific binding is characterized by binding isotherms that do not level off at a specific value of n due to a large number of such sites characterized by very weak (small) binding constants. More complete treatises concerned with the experimental aspect of the binding of ligands to macromolecules are available (cf. Ref. (3)).

17.6 BINDING OF CRO REPRESSOR PROTEIN TO DNA

As an example of a ligand binding study, we will consider some of the results obtained in the studies of the binding of the Cro repressor protein to DNA. DNA transcription is regulated by the binding of proteins to DNA. The Cro repressor is a protein that plays a regulatory role in λ phage. The binding of this protein to DNA has been examined by many different methods (8–10), and the three-dimensional structure of a Cro protein–DNA complex has been determined (11). The Cro protein binds as a dimer, and amino acids form DNA sequence-specific hydrogen bonds with exposed parts of DNA bases. Both the thermodynamics and the kinetics of Cro protein binding to DNA have been studied extensively.

The binding of Cro protein to DNA is difficult to study because the binding is very tight. In order to study the binding, a very sensitive filter binding assay was

TABLE 17-1. Binding Constants and Dissociation Rate Constants at 273 K for Cro Protein– λ DNA

Operator	K (M^{-1})	k_d (s^{-1})
Consensus	8.3×10^{11}	7.7×10^{-5}
OR3	5.0×10^{11}	1.7×10^{-4}
OR2	8.3×10^9	1.2×10^{-2}
OR1	1.2×10^{11}	1.7×10^{-3}
OL1	6.7×10^{10}	2.2×10^{-3}
OL2	3.7×10^{10}	2.6×10^{-3}
OL3	1.9×10^{10}	9.2×10^{-3}

Source: J. G. Kim, Y. Takeda, B. W. Matthews, and W. F. Anderson, *J. Mol. Biol.* **196**, 149 (1987).

used. Under certain conditions, linear duplex DNA passes through nitrocellulose filters, while a protein or protein–DNA complex is retained. Radioactive (^{32}P) DNA was prepared and mixed with Cro protein at very low concentrations (as low as <1 ng/ml). The reaction mixture was filtered within a few seconds, and from the radioactivity bound to the filter, the concentration of Cro protein–DNA could be calculated. Since the total concentrations of Cro protein and DNA are known, the equilibrium concentrations of all species can be calculated. The binding of Cro protein to various operator regions of λ phage DNA was studied. The operator is 17 base pairs long, but the DNAs investigated contained an extra 2 base pairs at each end. In addition, a consensus operator was constructed that had the consensus sequence of the six λ operator regions examined. The consensus sequence is identical to the OR3 operator, except for a few base pairs on the interior of the sequence.

The equilibrium association constants obtained are given in Table 17-1. The stoichiometry of the reaction was determined to be one to one. The association constants are very large, with the largest being for the consensus and OR3 operators. The OR3 operator was previously determined to be the preferred binding site for Cro protein. The binding of Cro operator to nonspecific DNA was also determined and found to be a linear function of the length of the DNA with a binding constant of $6.8 \times 10^5 M^{-1}$ per base pair, considerably weaker than the specific binding, as expected. The rate constants characterizing the reaction between the DNA operator sequences and the Cro protein were also determined. The dissociation rate constants are included in Table 17-1. The association rate constants that could be measured were all approximately $3 \times 10^8 M^{-1} s^{-1}$, the value expected for a diffusion controlled reaction. Every time a Cro protein and the operator DNA collide, a complex is formed.

The binding of Cro protein to long DNAs containing the operator region was also studied. The length of DNA varied from 73 to 2410 base pairs. For all except the shortest DNA, the equilibrium constant was about the same, $6.7 \times 10^{10} M^{-1}$. However, the association and dissociation rate constants increase as the length of DNA increased, leveling off at values of about $4.5 \times 10^9 M^{-1} s^{-1}$ and $1.7 \times 10^{-2} s^{-1}$, respectively. This very high second-order rate constant suggests that every time Cro protein collides with DNA, it binds very tightly. This result is puzzling as it would be expected that Cro protein would have to sample various parts of the DNA until it found the

operator and bound tightly. An explanation for these results is that Cro protein binds to the DNA on every collision and then diffuses rapidly along the DNA chain until it encounters the operator region. The rapid sliding of DNA binding protein along the DNA chain until it finds the correct place for a specific interaction appears to occur in other systems, but other mechanisms may also be operative.

Calorimetry was also carried out on the binding reaction (10). The results obtained indicate that the association of Cro protein with nonspecific DNA at 15°C is characterized by $\Delta H^\circ = 4.4 \text{ kcal/mol}$, $\Delta S^\circ = 49 \text{ cal/(mol K)}$, $\Delta G^\circ = -9.7 \text{ kcal/mol}$, and $\Delta C_p \approx 0$. The parameters obtained with the OR3 DNA are quite different, $\Delta H^\circ = 0.8 \text{ kcal/mol}$, $\Delta S^\circ = 59 \text{ cal/(mol K)}$, $\Delta G^\circ = -16.1 \text{ kcal/mol}$, and $\Delta C_p = -360 \text{ cal/(mol K)}$. In both cases, the favorable Gibbs energy change is entropy driven.

The specific molecular interactions between the DNA and the Cro protein have been probed by studying the binding of Cro protein to OR1 DNA, 21 base pairs as mentioned earlier, with systematic base substitutions along the entire DNA chain (9). Experiments were also carried out with point mutations in the protein of amino acids thought to be involved in the binding interaction. The changes in Gibbs energy were then interpreted as due to specific interactions between the protein and the DNA, and a structural map of the interactions was proposed. A high resolution crystal structure of a Cro protein–DNA complex determined some years later provided a definitive description of the molecular interactions (11). Many features proposed on the basis of the binding/mutation studies proved to be correct although some were not. As often stated in this text, molecular interpretations of thermodynamic studies must be viewed with caution. However, in retrospect, all of the mutation studies can be understood in terms of the crystal structure. A view of the overall complex is presented in Figure 17-5. Each Cro protein monomer consists of three α -helices and three β -strands. Only the helix–turn–helix portion of the protein makes direct contact with the DNA bases: the α_3 -helix is inserted into the major groove of DNA and its interactions with DNA bases account for the tight operator binding. A schematic map of the DNA base–Cro protein interactions is shown in Figure 17-6. Multiple hydrogen bonds are formed, although other types of interactions are also important. Both the protein and the DNA conformations are altered when the complex is formed. Many other interesting details of the molecular interactions can be inferred from the crystal structure but will not be considered here.

This study is an elegant demonstration of the type of information that can be obtained from thermodynamic, kinetic, and structural studies of ligand binding.

17.7 COOPERATIVITY IN LIGAND BINDING

We now return to the subject of macromolecules with multiple ligand binding sites. One of the most famous and best studied examples is hemoglobin. In Figure 17-7, the binding isotherm is shown for the binding of oxygen to myoglobin and hemoglobin. Hemoglobin contains four identical binding sites for oxygen. Myoglobin contains a single binding site for oxygen. The concentration of unbound oxygen is expressed

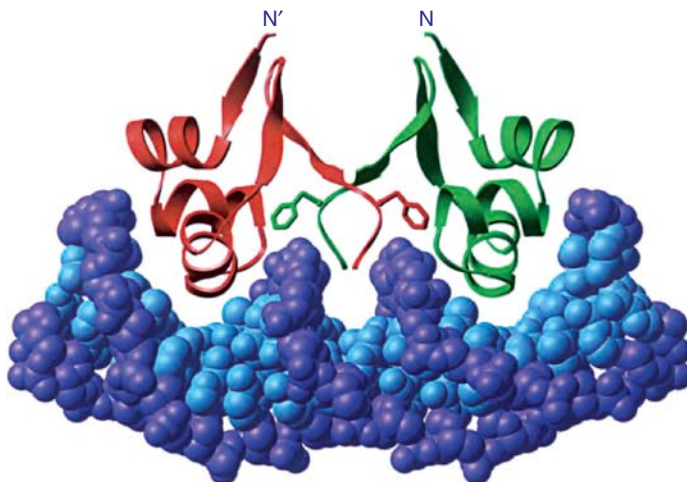


FIGURE 17-5. Structure of the overall complex between Cro protein and λ operator DNA. The direction of the view is parallel with the major grooves of DNA and parallel with the recognition helices. Reproduced from R. A. Albright and B. W. Matthews, "Crystal structure of lambda-Cro bound to a consensus operator" *J. Mol. Biol.* **280**, 137 (1998). Copyright 1998, with permission from Elsevier.

in pressure units of mmHg and rather than r ; the percent saturation (r/n) is shown so that the binding isotherms can be compared directly even though $n = 1$ for myoglobin and 4 for hemoglobin. The binding isotherm for myoglobin can be fit to Eq. (17-4) with $K = 0.23 \text{ (mmHg)}^{-1}$ [$K = [\text{MyO}_2]/([\text{My}][\text{O}_2])$, where My is myoglobin}. The data for hemoglobin clearly cannot be fit so simply, as the binding isotherm is not hyperbolic. Equation (17-21) can be used, however. Before doing so, it is useful to take into account the statistical effects associated with identical binding sites, that is, express Eq. (17-21) in terms of intrinsic binding constants rather than macroscopic binding constants by use of Eq. (17-17). With this substitution, Eq. (17-21) becomes

$$r = \frac{4K_1[\text{O}_2] + 12K_1K_2[\text{O}_2]^2 + 12K_1K_2K_3[\text{O}_2]^3 + 4K_1K_2K_3K_4[\text{O}_2]^4}{1 + 4K_1[\text{O}_2] + 6K_1K_2[\text{O}_2]^2 + 4K_1K_2K_3[\text{O}_2]^3 + K_1K_2K_3K_4[\text{O}_2]^4} \quad (17-25)$$

If this equation is used to fit the data obtained at pH 7.4, 25°C, 0.1M NaCl, the binding constants obtained are 0.024, 0.077, 0.083, and 7.1 (mmHg) $^{-1}$ (12). Note that the intrinsic binding constant becomes larger as each oxygen is bound. This is an example of cooperative binding. When an oxygen binds to hemoglobin, it increases the affinity of hemoglobin for the next oxygen. The result is a sigmoidal binding isotherm rather than a hyperbolic isotherm. This has important physiological consequences as it permits oxygen to be picked up and released over a very narrow range of oxygen pressure. Cooperative binding such as this is found frequently because it permits the biological activity to be regulated over a very narrow range of concentration.

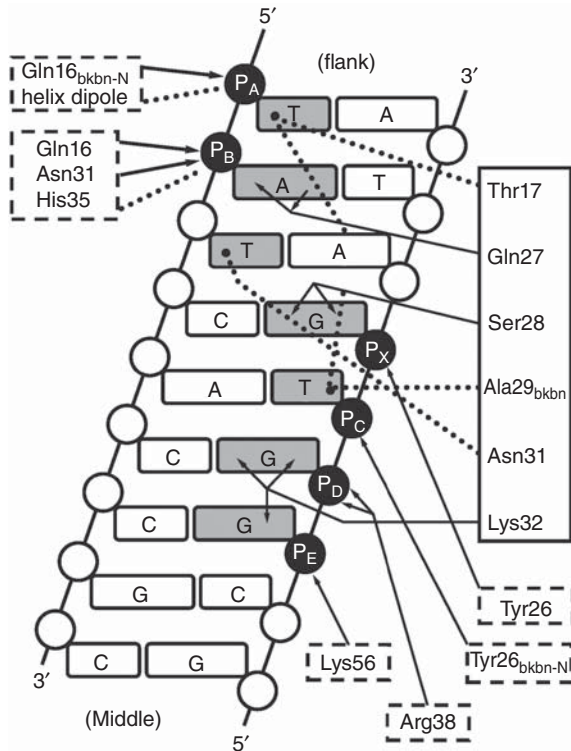


FIGURE 17-6. Schematic diagram of the interactions between the Cro protein and the λ operator DNA. The interactions for one of the polypeptide chains are shown, as the interactions with the other polypeptide chain of the dimeric Cro protein are symmetric. Hydrogen bonds are shown as continuous lines with arrows pointing from the donor to the acceptor. Broken lines are van der Waals contacts, bkn indicates a contact with the protein backbone, and the dotted lines are presumed electrostatic interactions. Reproduced from R. A. Albright and B. W. Matthews, crystal structure of lambda-Cro bound to a consensus operator, *J. Mol. Biol.* **280**, 137 (1998). Copyright 1998, with permission from Elsevier.

When the binding of oxygen by hemoglobin was first studied, the binding isotherm was fit by the empirical equation

$$\frac{r}{n} = \frac{[L]^\alpha / K^\alpha}{1 + [L]^\alpha / K^\alpha} \quad (17-26)$$

where the equilibrium constant is expressed as a dissociation constant and α is an empirical parameter called the Hill coefficient obtained from experiment by plotting $\ln[(r/n)/(1 - r/n)]$ versus $\ln[L]$ (13). {Note that $(r/n)/(1 - r/n) = [L]^\alpha / K^\alpha$.} As might be expected, such plots are linear over a limited range of ligand concentration. In the case of hemoglobin, α is about 2.5 and depends on the specific experimental

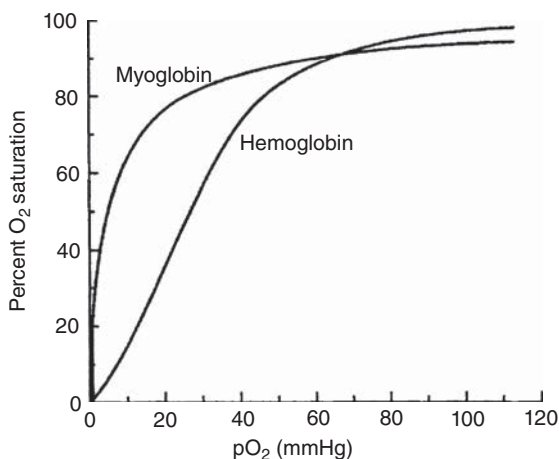
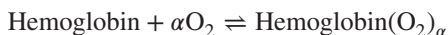


FIGURE 17-7. Binding of oxygen by myoglobin and hemoglobin at pH 7 and 38°C. The pressure of O₂ is a measure of the concentration of unbound O₂. Adapted from F. Daniels and R. A. Alberty, *Physical Chemistry*, 4th edition, John Wiley & Sons, Inc., New York, 1975. © 1975 John Wiley and Sons, Inc. Reprinted by permission of John Wiley & Sons, Inc.

conditions. The use of this equation does not have a physical meaning. It implicitly assumes that the binding equilibrium is



Obviously, the number of binding sites on hemoglobin must be an integer. Nevertheless, cooperative binding is frequently characterized by a Hill coefficient. The closer the Hill coefficient is to the actual number of binding sites, the more cooperative the binding.

It is also possible to have binding isotherms in which the binding constants decrease as successive ligands bind. This is usually termed *negative cooperativity* or *anticooperativity*, both oxymorons of a sort. It should be remembered that if macroscopic equilibrium binding constants are used, the binding constants decrease as each ligand is added to the macromolecule even if the sites are equivalent. This is the statistical effect embodied in Eq. (17-17). Anti- or negative cooperativity therefore means that the binding constants decrease more than the statistical effect expected for equivalent sites. An example of negative cooperativity is shown in Figure 17-8 for the binding of cytidine 3'-triphosphate (CTP) to the enzyme aspartate transcarbamoylase (14). This enzyme catalyzes the carbamylation of aspartic acid at a branch point in metabolism that eventually leads to the synthesis of pyrimidines. When the concentration of CTP becomes high, it inhibits aspartate transcarbamoylase and shuts down the metabolic pathway for its synthesis. As can be seen from the Scatchard plot presented, the data clearly do not conform to the straight line expected for independent equivalent binding sites. The total number of binding sites appears to be six, although extrapolation to this number is not precise. In this case,

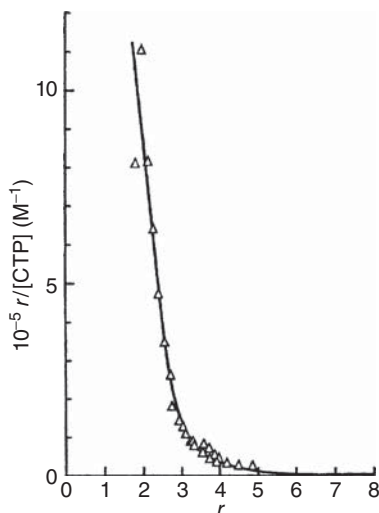


FIGURE 17-8. Scatchard plot for the binding of CTP to aspartate transcarbamoylase at 4°C and pH 7.3 in the presence of 2 mM carbamoyl phosphate and 10 mM succinate. Adapted from S. Matsumoto and G. G. Hammes, *Biochemistry* **12**, 1388 (1973). Copyright © 1973 American Chemical Society.

the data were fit to a model of two sets of three independent binding sites. The intrinsic binding constants obtained were $7.1 \times 10^5 \text{ M}^{-1}$ and $4.4 \times 10^3 \text{ M}^{-1}$. The data could be fit equally well to Eq. (17-21), but the data were not sufficient to define all six constants well. Aspartate transcarbamoylase is known to contain six identical binding sites for CTP and has three dimers of identical polypeptide chains. When CTP binds to one of the sites on a dimer, it weakens the binding for the second CTP. Interestingly, the binding of aspartate to this enzyme exhibits positive cooperativity, and the extent of the cooperativity is modulated by the binding of CTP, as will be discussed later.

The presence of positive or negative cooperativity can readily be diagnosed from the binding isotherm. The shapes of the various plots commonly used are shown in Figure 17-9 for no cooperativity, positive cooperativity, and negative cooperativity. The type of cooperativity occurring can readily be diagnosed from these plots. It is particularly obvious in Scatchard plots as no cooperativity gives a straight line, a maximum in the plot is observed for positive cooperativity, and a concave curve is found for negative cooperativity. For some studies, $\log[L]$ is used in order that a wide range of concentrations can be represented in a single plot (Fig. 17-9b).

When positive and negative cooperativities are observed, one must determine if this is due to a preexisting difference between the binding sites, or if the binding of one ligand alters the binding affinity of the remaining sites for the ligand. Usually, but not always, the latter is true, as for the cases cited earlier. The possibility also exists that both positive and negative cooperativity could occur in a single binding isotherm.

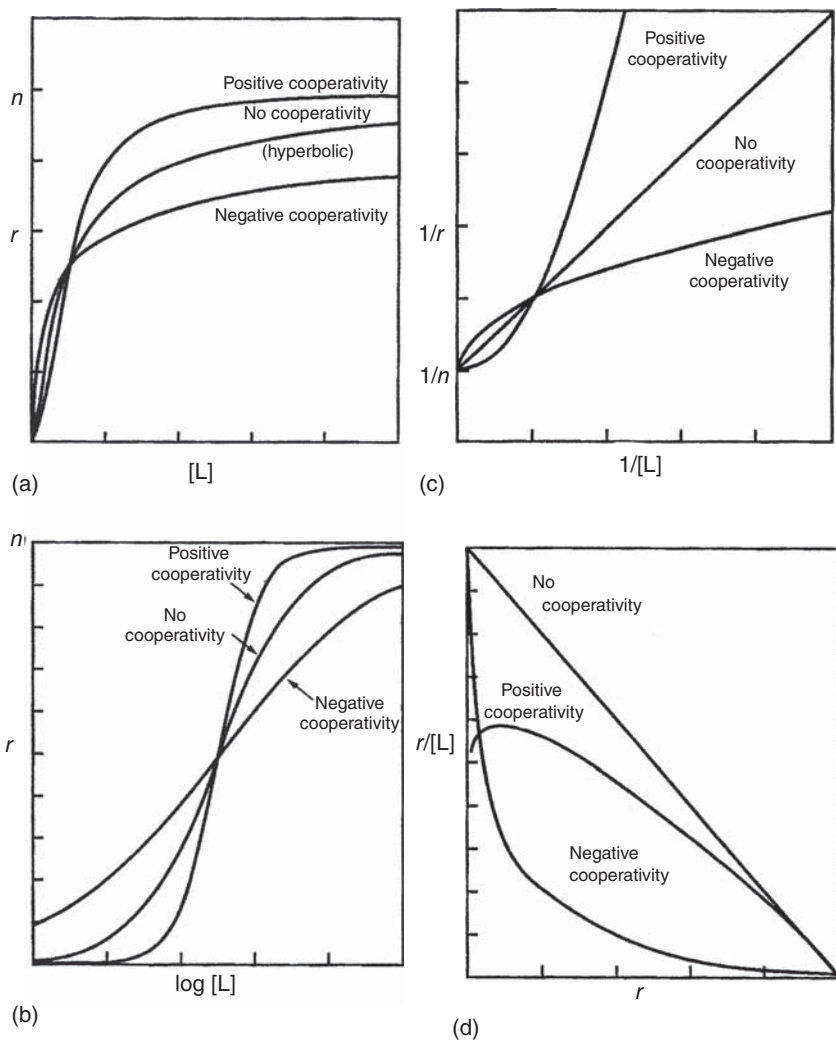


FIGURE 17-9. Schematic representations of equilibrium binding data demonstrating no cooperativity, positive cooperativity, and negative cooperativity. In these figures, r is the moles of ligand bound per mole of protein and $[L]$ is the concentration of unbound ligand.

17.8 MODELS FOR COOPERATIVITY

Thus far, we have been concerned with fitting data to binding isotherms and developing criteria with regard to whether the binding is cooperative. However, the ultimate goal is to relate the experimental data to molecular structure. This requires the development of theoretical models that relate structure to ligand binding isotherms. Two limiting models have been developed to explain cooperative ligand binding to proteins. Both models are based on the general hypothesis that cooperativity is the

result of alterations in the interactions between polypeptide chains through conformational changes in the macromolecule. One of the models assumes a concerted or global conformational change, whereas the other assumes a sequential change in the conformation of each polypeptide chain or chains that contain a ligand binding site.

We first consider the concerted model that has been developed by Monod, Wyman, and Changeux (15). This model (MWC) is based on three assumptions: (i) the protein consists of two or more identical subunits, each containing a site for the ligand; (ii) at least two conformational states, usually designated as R and T states, are in equilibrium and differ in their affinities for the ligand; and (iii) the conformational changes of all subunits occur in a concerted manner (conservation of structural symmetry). A schematic illustration of the MWC model for a four-subunit protein is shown in Figure 17-10. A sigmoidal binding isotherm can be generated from this model in the following way. In the absence of ligand, the protein exists largely in the T state (the square conformation), but the substrate binds preferentially to the R state (the circular conformation). When the ligand binds, it shifts the protein from the T to the R state. Thus, at low ligand concentrations, the protein is primarily in the T state, whereas at high ligand concentration, the protein is largely in the R state. This shift in equilibria can give rise to a sigmoidal binding isotherm, or positive cooperativity. This model cannot, however, explain anti(negative) cooperativity.

The quantitative development of this model is complex, although not difficult, and will only be given in outline form here. The scheme for ligand binding can be represented as

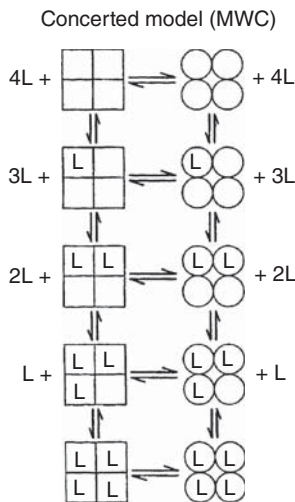
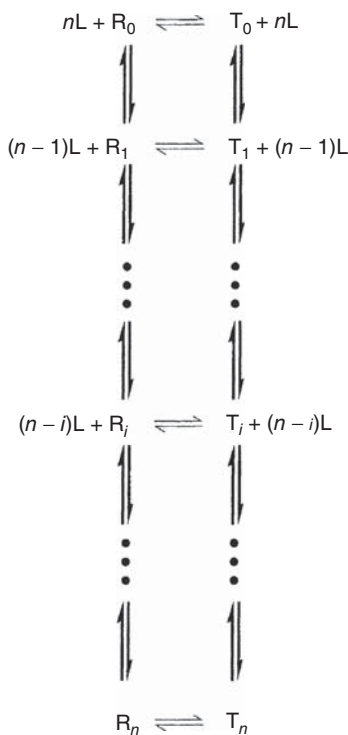


FIGURE 17-10. Schematic representation of the Monod–Wyman–Changeux model for a four-subunit protein. The squares and circles designate different subunit conformations of the protein, and L is the ligand.



where R_0 and T_0 are the two different conformational states in the absence of ligand, and R_i and T_i designate their complexes with i molecules of L . Three constants are needed to specify the equilibrium binding isotherms: the intrinsic dissociation constants for ligand binding to the R and T states, and the equilibrium constant for the ratio of the R_0 to T_0 states. These constants can be written as

$$\begin{aligned}
 L_0 &= [T_0]/[R_0] \\
 K_R &= [(n-i+1)/i][R_{i-1}][L]/[R_i] \\
 K_T &= [(n-i+1)/i][T_{i-1}][L]/[T_i]
 \end{aligned} \tag{17-28}$$

The fraction of sites occupied by the ligand, Y , can be expressed as

$$\begin{aligned}
 Y &= \frac{r}{n} = \frac{([R_1] + 2[R_2] + \cdots + n[R_n]) + ([T_1] + 2[T_2] + \cdots + n[T_n])}{n\{([R_0] + [R_1] + \cdots + [R_n]) + ([T_0] + [T_1] + \cdots + [T_n])\}} \\
 Y &= \frac{L_0 c \alpha (1 + c \alpha)^{n-1} + \alpha (1 + \alpha)^{n-1}}{L_0 (1 + c \alpha)^n + (1 + \alpha)^n}
 \end{aligned} \tag{17-29}$$

where $\alpha = [L]/K_R$ and $c = K_R/K_T$. The transformation of the first part of Eq. (17-29) to the second part requires the use of the binomial theorem and will not be detailed here.

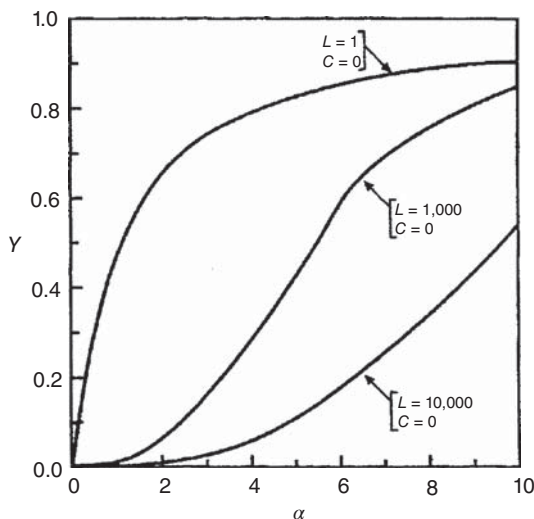


FIGURE 17-11. Plot of the fraction of sites bound by ligand, Y , versus $\alpha (= [L]/K_R)$ for various values of c and L according to Eq. (17-29). A sigmoidal binding isotherm is generated when L is large and c is small.

The nature of the binding isotherm depends on the values of L_0 and c . A hyperbolic isotherm is obtained when the ligand binds equally well to both conformations, $c = 1$, and when both L_0 and c are either very large or very small. However, when L_0 is large and c is small, sigmoidal binding isotherms occur as illustrated in Figure 17-11. When c is very small, Eq. (17-29) becomes

$$Y = \frac{\alpha(1 + \alpha)^{n-1}}{L_0 + (1 + \alpha)^n} \quad (17-30)$$

With this limiting case, it can easily be seen that a hyperbolic isotherm is found when L_0 is small, whereas a sigmoidal isotherm is predicted when L_0 is large.

The basic assumptions of an alternative model developed by Koshland, Nemethy, and Filmer (KNF) are the following (16): (i) two conformational states, A and B, are available to each subunit; (ii) only the subunit to which the ligand is bound changes its conformation; and (iii) the ligand-induced conformational change in one subunit alters its interactions with the neighboring subunits. The strength of the subunit interactions may be increased, may be decreased, or may stay the same. The result of this change in subunit interactions is that the binding of the next ligand can be weaker, stronger, or the same as the binding of the previous ligand. Clearly, this model can produce either positive or negative cooperativity—or a hyperbolic binding isotherm. This sequential model is shown schematically in Figure 17-12 for a four-subunit protein in a square configuration.

Calculation of the binding isotherm for the KNF model is complex and will not be presented here. The basic parameters are the intrinsic binding constant, an

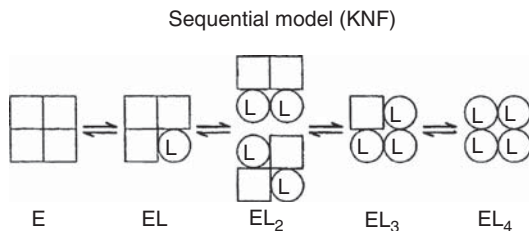


FIGURE 17-12. Schematic representation of the Koshland–Nemethy–Filmer model for a four-subunit (square) protein. The squares and circles designate different subunit conformations, and L is the ligand. Note that two structures are shown for the intermediate with two ligands bound as the subunit interactions are different for these two species.

equilibrium constant characterizing the conformational change that occurs, and constants characterizing the subunit interactions, AB and BB. (The AA interaction is taken as the reference state, so it does not appear in the calculation.) The binding isotherm that results is identical in form to Eq. (17-21), except that the appropriate statistical corrections are included so that the intrinsic binding constant appears. The effective binding constant multiplying each successive power of L may increase, decrease, or stay the same, depending on the nature of the subunit interactions, so that positive, negative, or mixed cooperativity is possible.

The MWC and KNF models are limiting cases of a more general scheme shown in Figure 17-13. This figure illustrates a general mechanism involving a tetrameric protein and only two conformational states for each subunit. The real situation is somewhat more complex, as the permutations of the ligand among the subunits for a given conformational state are not shown. The extreme right- and left-hand columns enclosed by dashed lines represent the MWC mechanism, whereas the diagonal, enclosed by dotted lines, represents the KNF model. Thus, these two models are limiting cases of an even more complex scheme. As might be suspected, distinguishing among these models when positive cooperativity occurs is not a simple task.

17.9 KINETIC STUDIES OF COOPERATIVE BINDING

Thus far, we have confined the discussion to ligand binding at equilibrium. The binding process and the cooperativity that may occur play an important role in many biological processes, as evidenced by hemoglobin, membrane receptor binding, and operator binding to DNA. The role of cooperativity in regulating such reactions has been well established. However, many enzyme reactions are also regulated through cooperative binding processes. In such cases, the rate of the enzymatic reaction is usually measured, often the initial velocity, and it is observed that a plot of the initial velocity versus the substrate concentration exhibits cooperativity. An example already discussed is aspartate transcarbamoylase: a plot of the initial velocity versus the aspartate concentration is sigmoidal so that small changes in the concentration of aspartate can cause significant changes in the rate of the enzymatic reaction.

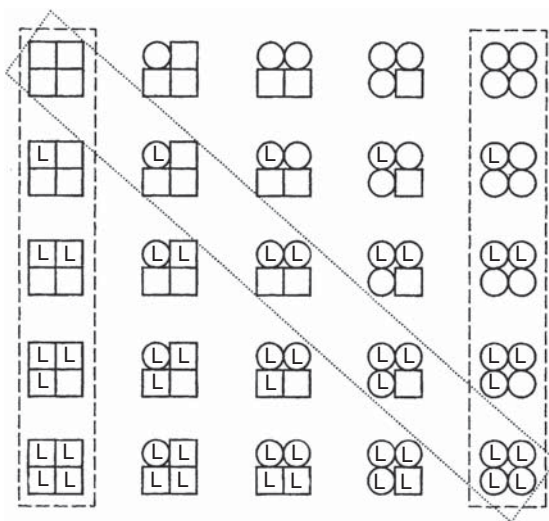


FIGURE 17-13. Schematic illustration of a general allosteric model for the binding of a ligand to a four-subunit protein. The squares and circles designate different subunit conformations. The portion enclosed by dashed lines is the Monod–Wyman–Changeux model, whereas that enclosed by dotted lines is the Koshland–Nemethy–Filmer model. For the sake of simplicity, the permutations of the ligand among the subunits and the free ligand are omitted.

For many enzymes, plots of the initial velocity versus the substrate concentration are essentially identical in form with the binding isotherm. This suggests that the binding steps prior to the rate-determining step are rapid and reversible and that the turnover numbers (catalytic efficiency) for all of the binding sites are identical. If this is the situation, the initial velocity, v , can readily be related to the binding isotherm:

$$v = \frac{V_m r}{n} \quad (17-31)$$

where V_m is the maximum velocity of the enzyme reaction expressed as the product of the molar concentration of enzyme and the turnover number. This simple analysis seems to be sufficient for many cases, and if it is, the treatment of the rate data parallels what was done for equilibrium binding, including the methods of plotting and fitting the data. As might be expected, both positive and negative cooperativity are observed.

The interpretation of kinetic data can, however, be more complex. The turnover numbers for different sites could be different. For example, in terms of the MWC model, the R and T forms might have different catalytic activities although usually the assumption is made that only the R form is enzymatically active. For the KNF model, each of the ligand binding sites might have a different catalytic activity. In fact, these complications are rarely included, or justified, in the data analysis. Finally, it should be noted that apparent cooperativity in the rate of an enzymatic reaction can arise purely from special kinetic situations and may have nothing to do with cooperative binding of substrate. A few such situations have been well documented,

but we will not consider such complications further. This is just a reminder that kinetic measurements are not a substitute for equilibrium binding studies. They may provide similar information in some cases, but they are inherently more difficult to interpret. Of course, conversely, kinetic studies can provide dynamic information that cannot be obtained from equilibrium measurements.

17.10 ALLOSTERISM

Thus far, we have considered only cases where a single ligand binds to a macromolecule, and important biological control can occur through cooperative interactions for a single ligand. These are *homotropic* interactions. However, in many instances regulation occurs through the binding of a second ligand that influences the binding of the first ligand. These are called *heterotropic* interactions. Regulatory control by reaction of a second ligand is termed *allosterism*, and the second binding site is called an *allosteric* site. Two specific examples will be discussed to illustrate the principles involved: hemoglobin and aspartate transcarbamoylase.

The association of oxygen with hemoglobin is strongly pH dependent as shown in Figure 17-14. Note that the oxygenation isotherm becomes more sigmoidal as the pH is lowered. Thus, the binding of protons to hemoglobin clearly affects oxygen binding. In fact, the addition of protons decreases the amount of oxygen bound, and vice versa. This reciprocal relationship is called the Bohr effect. The effect of proton binding is typical allosteric regulation, in this case of oxygen binding. How is this accommodated in the models we have discussed previously? For the MWC model, allosteric effectors are assumed to bind selectively to the R or T conformation. An inhibitor, I, such as the proton in the case under consideration, would bind selectively to the T conformation. This shifts the equilibrium from the R to the T state, which effectively changes the equilibrium constant L_0 to

$$L'_0 = L_0 \left(1 + \frac{[I]}{K_I} \right)^n \quad (17-32)$$

where K_I is the intrinsic dissociation constant for the binding of inhibitor, and it has been assumed that a single proton binds to each of the four subunits. Insertion of this relationship into Eq. (17-29) can quantitatively account for the Bohr effect. The KNF model can also explain the observations by assuming that binding of the inhibitor alters the subunit interactions and conformational changes, effectively altering the binding constants for oxygen as successive ligands are bound. Organic phosphates are also strong effectors of oxygen binding and bind preferentially to deoxyhemoglobin.

An activator, A, can enhance ligand binding by binding selectively to the R conformation. This effectively changes the equilibrium constant L_0 to

$$L'_0 = \frac{L_0}{(1 + [A]/K_A)^n} \quad (17-33)$$

where K_A is the intrinsic dissociation constant for the binding of the activator.

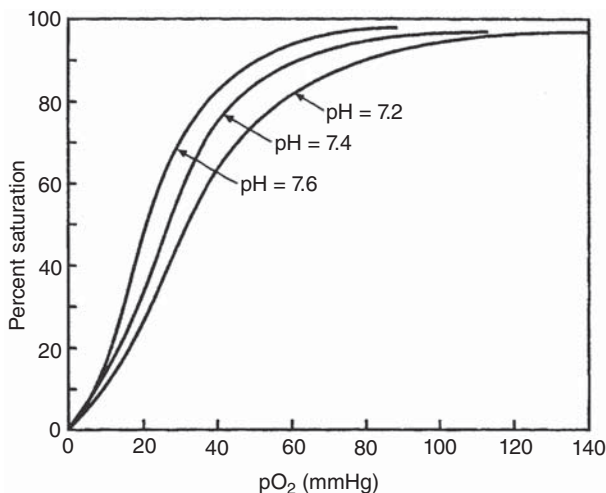


FIGURE 17-14. The effect of pH on the oxygenation of hemoglobin (the Bohr effect). The percent of saturation by oxygen is plotted versus the pressure of O₂. Adapted from R. E. Benesch and R. Benesch, *The mechanism of interaction of red cell organic phosphates with hemoglobin*, *Adv. Protein Chem.* **28**, 211 (1974). Copyright 1974, with permission from Elsevier.

Which of the models for allosterism best accommodates the known data for hemoglobin? The nature of the conformational change occurring when oxygen binds is known from structural studies of deoxy- and oxyhemoglobin. The hemoglobin studied contains four polypeptide chains, two α chains and two β chains. The α and β chains are similar but not identical. When oxygen binds, salt bridges are broken between the polypeptide chains, and all four chains rotate slightly to accommodate the movement of iron into the plane of the heme. The iron moves only a few tenths of an angstrom, and two of the polypeptides move about 7 Å closer (17). These structural changes and the binding data fit quite well to the MWC model. However, there is evidence that in addition to this major conformational change, sequential conformational changes occur in the subunits. Therefore, it is likely the case that both models are needed to accommodate the data, and multiple conformational changes occur. The structure of hemoglobin and the proposed major concerted conformational change are shown schematically in Figure 9-9.

As previously indicated, aspartate transcarbamoylase is a key enzyme in the pathway for pyrimidine biosynthesis. It is subject to inhibition by CTP and to activation by ATP. The allosteric binding of CTP makes the dependence of the rate on aspartate concentration more sigmoidal, whereas binding of the activator, ATP, makes the curve less sigmoidal (18). This is a prototype for feedback inhibition in metabolism and is shown schematically in Figure 17-15. An additional wrinkle in the regulatory process is that the binding of both CTP and ATP to the enzyme displays negative cooperativity, and these two ligands compete for the same binding site. The effect

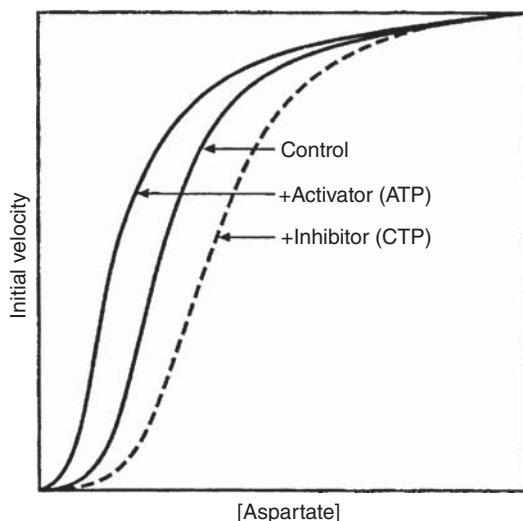


FIGURE 17-15. Schematic representation of the dependence of the initial velocity of the reaction catalyzed by aspartate transcarbamoylase on the concentration of aspartate, a substrate. The effects of an allosteric inhibitor, CTP, and an allosteric activator, ATP, are shown.

of ATP and CTP on the binding of aspartate can readily be accommodated by the MWC model with the assumption that CTP binds selectively to the T conformation and ATP to the R conformation. This assumes that aspartate binds selectively to the R conformation. This model, however, cannot accommodate the negative cooperativity observed in the binding of CTP and ATP. The KNF model can also accommodate these results through alterations in the interactions between subunits.

The structure of aspartate transcarbamoylase is known. It consists of two trimers of catalytic sites, with three dimers of regulatory sites at the interface of the trimers (19). A threefold rotational axis is present, and it has been shown that conversion of the putative R to T form involves rotation around this axis and alteration of the interactions between the regulatory and catalytic subunits. The MWC model accommodates much of the data, but the data also require sequential conformational changes in the subunits. Thus, the conclusion is similar to that for hemoglobin. A major conformational change consistent with R and T conformations appears to occur, but additional conformational changes more localized and sequential in nature also occur. The regulatory process, therefore, seems to require multiple conformations and the interplay between global and local conformational changes. The use of multiple conformational changes enhances the versatility and sensitivity of the regulatory process. The structure of aspartate transcarbamoylase and the nature of the concerted conformational change are shown in Figure 17-16. The top of the figure is the T state with the catalytic trimers in blue and white at the top and bottom of the structure. Two of the regulatory dimers are at the sides of the structure in yellow. The third dimer is at the back of the structure. The bottom structure is the R state. The movement of the subunits with respect to each other can be clearly seen.

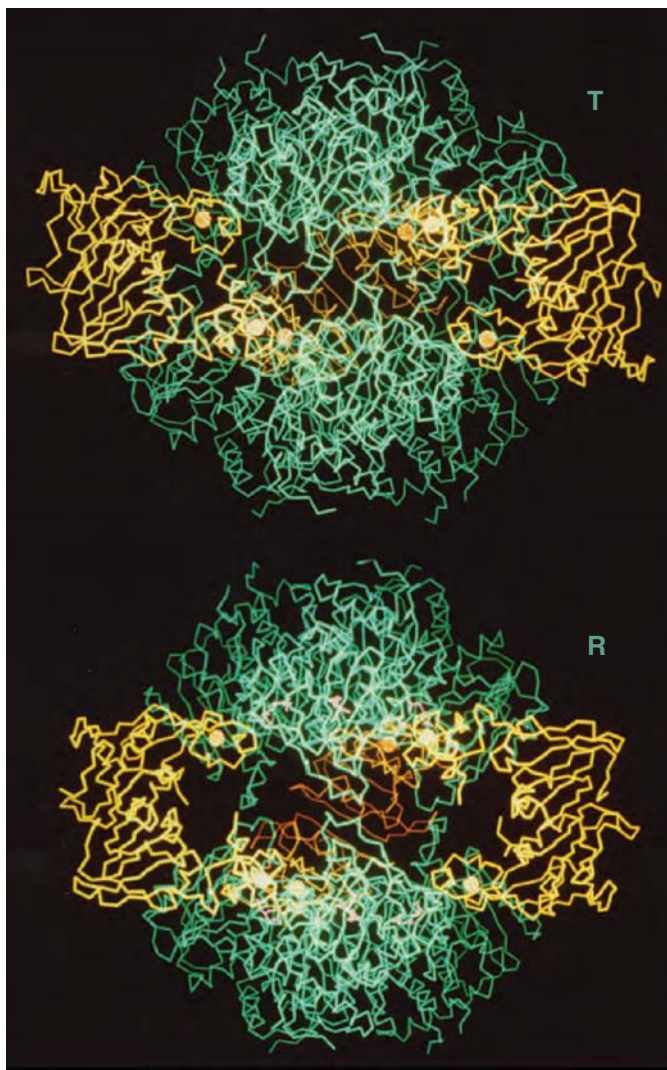


FIGURE 17-16. Ca backbone structures of the R (lower) and T (upper) states of aspartate transcarbamoylase. The two catalytic trimers are in green and are closer together in the T state. The three regulatory dimers (equatorial) are in yellow. One of the regulatory dimers is behind the large central cavity in this view. Adapted from W. N. Lipscomb, *Adv. Enzymol.* 68, 67 (1994). Copyright 1994 John Wiley & Sons, Inc. Reprinted by permission of John Wiley & Sons, Inc.

This concludes our discussion of ligand binding in biology. We have developed the theoretical and experimental frameworks and have provided several examples of applications to biological systems. This, and further reading of the literature, should permit the interested reader to develop specific applications as needed.

REFERENCES

1. G. G. Hammes, *Enzyme Catalysis and Regulation*, Academic Press, New York, 1982.
2. C. R. Cantor and P. R. Schimmel, *Biophysical Chemistry, Parts I, II, and III*, Freeman, San Francisco, CA, 1980.
3. I. M. Klotz, *Ligand-Receptor Complexes*, John Wiley & Sons, New York, 1997.
4. G. Scatchard, *Ann. N. Y. Acad. Sci.* 51, 660 (1949).
5. A. O. Pedersen, B. Hust, S. Andersen, F. Nielsen, and R. Brodersen, *Eur. J. Biochem.* 154, 545 (1986).
6. R. Brodersen, B. Honoré, A. O. Pedersen, and I. M. Klotz, *Trends Pharm. Sci.* 9, 252 (1988).
7. D. G. Anderson, G. G. Hammes, and F. G. Walz, Jr., *Biochemistry* 7, 1637 (1968).
8. J. G. Kim, Y. Takeda, B. W. Matthews, and W. F. Anderson, *J. Mol. Biol.* 196, 149 (1987).
9. Y. Takeda, A. Sarai, and V. M. Rivera, *Proc. Natl. Acad. Sci. USA* 86, 439 (1989).
10. Y. Takeda, P. D. Ross, and C. P. Mudd, *Proc. Natl. Acad. Sci. USA* 89, 8180 (1992).
11. R. A. Albright and B. W. Matthews, *J. Mol. Biol.* 280, 137 (1998).
12. I. Tyuma, K. Imai, and K. Shimizu, *Biochemistry* 12, 1491 (1973).
13. A. V. Hill, *J. Physiol.* 40, 4 (1910).
14. S. Matsumoto and G. G. Hammes, *Biochemistry* 12, 1388 (1973).
15. J. Monod, J. Wyman, and J.-P. Changeux, *J. Mol. Biol.* 12, 88 (1965).
16. D. E. Koshland, Jr., G. Nemethy, and G. Filmer, *Biochemistry* 5, 365 (1966).
17. J. M. Friedman, *Science* 228, 1273 (1985).
18. J. C. Gerhart and A. B. Pardee, *J. Biol. Chem.* 237, 819 (1962).
19. W. N. Lipscomb, *Adv. Enzymol.* 68, 67 (1994).

PROBLEMS

17-1. The following data were obtained for the binding of ADP to an ATPase.

How many binding sites are present per mole of enzyme and what is the intrinsic binding constant? The concentration in the table is unbound ADP. (Nonspecific binding occurs, which is not uncommon. You will have to decide how to handle this problem.)

R	[ADP] (μM)
0.500	0.719
0.694	1.23
1.06	2.36
1.22	3.30
1.32	4.55
1.63	8.81
1.69	16.32
1.91	27.15
2.21	40.62
2.15	79.75
2.40	115.5

- 17-2.** Some typical equilibrium dialysis data for the binding of a ligand to a macromolecule are given below. The total concentration of the macromolecule is $0.500\ \mu\text{M}$.

Determine the number of binding sites on the macromolecule and the intrinsic binding constant.

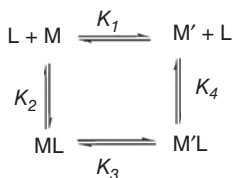
Total ligand concentration (μM)	
Side without macromolecule	Side with macromolecule
0.158	0.436
0.395	0.960
0.890	1.83
2.37	3.78
4.00	5.60
6.18	7.91
8.12	9.90

- 17-3.** The following initial velocities, v , were measured for the carbamylation of aspartic acid by carbamoyl phosphate as catalyzed by the enzyme aspartate transcarbamoylase. The concentration of carbamoyl phosphate was $1\ \text{mM}$ and the concentration of aspartate was varied.

v (arbitrary units)	Aspartate (mM)
0.90	2.0
1.60	3.0
2.25	4.0
3.20	5.0
3.65	6.0
4.70	8.0
5.05	10.0
5.25	12.0
5.80	15.0
6.00	17.0
6.05	20.0

- a.** Assume that the initial velocity is related to r by Eq. (17-31) and construct a Hill plot of the data. The slope provides a lower bound to the number of aspartate binding sites.

- b. In fact, six aspartate binding sites are present per mole of protein. Using this information and the data, make a table of r and the corresponding aspartate concentration. Make a plot of $r/[\text{aspartate}]$ versus r .
- c. What type of cooperativity is occurring? Which of the two limiting models discussed (MWC and KNF) is consistent with the data?
- 17-4. Consider a macromolecule that has two different conformations, M and M'. The two conformations bind a single ligand, L, per molecule, but the binding equilibrium is different for each conformation. This can be represented as



where K_i are the equilibrium constants for the individual reactions. Calculate the binding isotherm, r , for this macromolecule in terms of the equilibrium constants and the concentration of the unbound ligand. What type of cooperativity, if any, is displayed by this system? How many of the individual equilibrium constants can be determined from the binding isotherm? What relationship, if any, exists between the four constants?

Hydrodynamics of Macromolecules

18.1 INTRODUCTION

Hydrodynamics is the science of the movement and interactions of particles in fluids, both in the absence and in the presence of external forces such as centrifugal fields (centrifugation) and electrical fields (electrophoresis). Hydrodynamic measurements are used frequently to characterize the behavior of macromolecules in solution and can provide important molecular information. In this chapter, we consider some of the fundamental aspects of hydrodynamics in relatively simple terms and analyze some of the methods used to obtain molecular information about macromolecules. The mathematical details are not presented: more detailed discussions are available (cf. Refs. 1–4).

18.2 FRICTIONAL COEFFICIENT

If a molecule moves through a solution, it is subjected to a frictional drag because of interactions with the solvent. If the movement is slow enough so as not to cause turbulence, the frictional drag is directly proportional to the velocity of the molecule, and the proportionality constant is called the *frictional coefficient*, f . The actual calculation of this frictional coefficient is quite complex, and we present only the results. If the molecule is assumed to be a sphere of radius R_S , the frictional coefficient is

$$f = 6\pi\eta R_S \quad (18-1)$$

where η is the viscosity of the liquid. This relationship is for the translational frictional coefficient. The frictional coefficient for rotational motion can also be calculated:

$$f_{\text{rot}} = 6\eta V \quad (18-2)$$

where V is the volume of the sphere, $(4/3)\pi R_S^3$. Two immediate questions arise with regard to these definitions, namely, what is the “volume,” and what if the macromolecule cannot be approximated as a sphere.

Defining the volume turns out to be more difficult than might be imagined. What can be measured precisely is the partial specific volume of the solute, $\bar{V}_2$. In thermodynamic terms, the partial specific volume of a solute is the change in solution volume when a small increment of solute is added to the solution, extrapolated to infinite dilution. (Partial molar quantities are discussed in Chapter 4.) In practice, the volume of the solution, V_{total} , is determined as a function of the weight concentration of the solute, g . At a given constant weight, temperature, and pressure, $\bar{V}_2 = dV_{\text{total}}/dg$, and the observed value of $\bar{V}_2$ at various values of g is extrapolated to zero concentration to eliminate the effect of interactions between macromolecules. In practice, $\bar{V}_2$ is rarely measured directly because of the large amount of solute required. Instead, it is estimated from the amino acid composition of the protein. The partial specific volumes of the amino acids have been carefully measured so that the partial specific volume of the protein can be calculated by summing the partial specific volumes of the amino acids on a weight basis. For proteins, $\bar{V}_2$ ranges from 0.69 to 0.76 cm³/g, and a value of 0.73 cm³/g is used in the absence of amino acid data. Nucleic acids are more dense, and the value of $\bar{V}_2$ is more sensitive to temperature and solvent, particularly the specific counterion, because the strong electrostatic interactions between the polyphosphate chain and metals alter the hydration of the macromolecule. A value of about 0.5 cm³/g is obtained if the counterion is Na⁺.

Since $\bar{V}_2$ is calculated on a weight basis, the observed volume of the molecule is $(M/N_0)\bar{V}_2$, where M is the molecular weight and N_0 is Avogadro's number. This determination, however, does not take into account the hydration sphere associated with the molecule. The hydration of the macromolecule can be included by adding the volume of hydration $\delta\bar{V}_1$, where δ can be estimated by experimental measurements not considered here and $\bar{V}_1$ is the specific volume of the solvent. Thus, the hydrated volume of the molecule, V_h , is

$$V_h = \left(\frac{M}{N_0}\right)(\bar{V}_2 + \delta\bar{V}_1) \quad (18-3)$$

This should be regarded as an empirical estimate of the hydrated volume. The first term is often considered the anhydrous volume of the macromolecule and the second term the water of hydration. However, this is an oversimplification: $\bar{V}_2$ is also dependent on the solvent. As further validation of this approximate model, the thickness of the water layer around the macromolecule can be calculated to be about one layer. Because of the uncertainty in the extent of hydration of a macromolecule and its dependence on the details of the molecular surfaces, the hydrodynamic radius is sometimes calculated from $\bar{V}_2$, that is, not explicitly including hydration. This radius is often called the hydrodynamic or Stokes' radius. George Stokes (1819–1903) was one of the pioneers in the field of hydrodynamics, and Eq. (18-1) is often called Stokes' law.

Most macromolecules are not spheres: this is obvious for nucleic acids and even in the case of proteins a better representation would be some type of irregularly shaped rigid body. Nevertheless, a sphere is the most common shape analyzed because of the relative simplicity of the mathematics. In fact, even when molecules are known to be

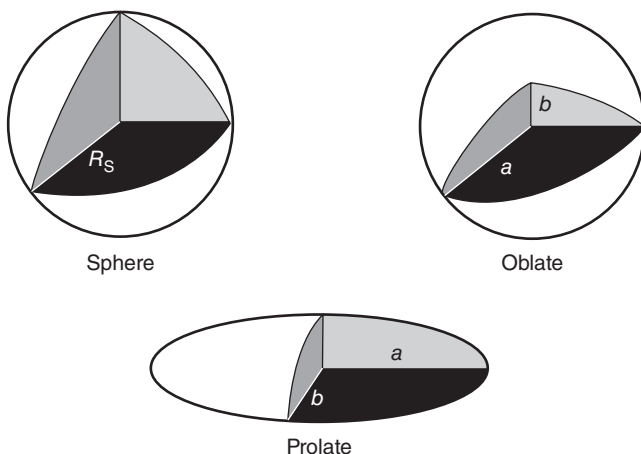


FIGURE 18-1. The most common hydrodynamic shapes used as models for molecules: the sphere and oblate and prolate ellipsoids. The cross sections are shown to illustrate the shapes. The a and b axes are defined as the length and width of the entire ellipsoid. Adapted from Ref. 1.

nonspherical, an “apparent” Stokes’ radius is often calculated with Eq. (18-1). Two other shapes have been analyzed in detail, prolate and oblate ellipsoids of revolution. These shapes are shown in Figure 18-1, along with a sphere. A prolate ellipsoid is cigar shaped and is generated by rotation about its long axis. The two parameters characterizing it are the long axis, a , and the short axis, b . Obviously for a sphere, $a = b$. An oblate ellipsoid is generated by rotation around its short axis, b , and resembles a disk in shape. As might be guessed, these ellipsoids of revolution have larger translational frictional coefficients than spheres.

Tables of the ratio of the frictional coefficient of the ellipsoid to that of a sphere are available as a function of the axial ratio, a/b (prolate ellipsoids) or b/a (oblate ellipsoids). An abbreviated compilation is given in Table 18-1. This frictional coefficient ratio, f/f_{sphere} , varies from 1 to 6 for a prolate ellipsoid and from 1 to 4 for an oblate ellipsoid when the axial ratio goes from 1 to 200. As will be discussed later, hydrodynamic measurements can be used to infer information about the shape of macromolecules.

Rotational frictional coefficients are more complex because rotation can occur around either the long or short axis of an ellipsoid, so that two frictional coefficients are needed for each ellipsoid. This case will not be dealt with further except to note the following. For oblate ellipsoids, the rotational friction coefficient is similar about both axes and larger than that of the sphere. For prolate ellipsoids, the two friction coefficients are quite different, with rotation around the long axis having a smaller friction coefficient than the equivalent sphere and rotation around the short axis having a larger friction coefficient. This is qualitatively what might be expected.

Frictional coefficients for more complex shapes have been calculated but are not considered here.

TABLE 18-1. Frictional Coefficient Ratios and Intrinsic Viscosity Shape Factors for Ellipsoids of Revolution

Axial ratio	Prolate		Oblate	
	ν	f/f_{sphere}	ν	f/f_{sphere}
1	2.500	1.000	2.500	1.000
5	5.806	1.250	4.708	1.224
10	13.634	1.543	8.043	1.458
20	38.53	1.996	14.80	1.782
50	176.81	2.946	35.16	2.375
100	593.7	4.067	69.10	2.974
200	2052.9	5.708	137.01	3.735

Data from Ref. 1 and H. A. Scheraga, *Protein Structure*, Academic Press, New York, 1961.

18.3 DIFFUSION

The motion of molecules in solution in the absence of external forces is due to diffusion. In a solution at equilibrium, the motion of molecules is random and caused by thermal energy: as the temperature goes up, the speed of the molecules increases. Diffusion cannot be easily observed in solutions at equilibrium: sophisticated light scattering techniques are required. The diffusion of macromolecules can be studied more directly by creating a boundary between a buffer and a solution of the macromolecule. As time passes, the boundary spreads due to diffusion. The movement of molecules with time is governed by Fick's laws. The first law states that when a concentration gradient exists, dc/dx , the number of molecules, n , passing through an area, A , per unit time, dn/dt , is proportional to the concentration gradient, with the proportionality constant defining the *diffusion constant*, D . This can be written in one dimension, x , as

$$\frac{dn}{dt} = -DA \left(\frac{dc}{dx} \right) \quad (18-4)$$

Strictly speaking, these should be partial derivatives, that is, the derivative with respect to time assumes that x is constant and the derivative with respect to x assumes that the time is constant. The minus sign arises because the flow of molecules is opposite in sign to the concentration gradient, and D is defined as a positive number. From dimensional analysis, D can be seen to have the dimensions of cm^2/s .

Fick's first law is easy to understand on an intuitive basis. It states that the net rate of flow of molecules is proportional to the concentration gradient, dc/dx : the larger the concentration gradient, the faster the flow. If there is no concentration gradient ($dc/dx = 0$), no net flow occurs. Fick's second law is not as apparent. It is based on the first law and the conservation of mass and relates the time dependence of the concentration to the concentration gradient. Often it is more convenient to measure the concentration rather than the flux (dn/dt). Fick's second law in one dimension can be written as

$$\frac{dc}{dt} = D \left(\frac{d^2c}{dx^2} \right) \quad (18-5)$$

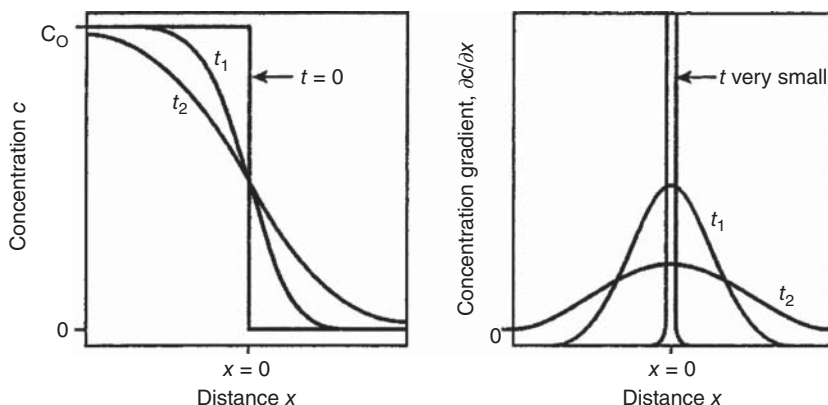


FIGURE 18-2. The progress of a diffusion experiment that begins with a sharp boundary. The concentration is shown at three different times on the left and the concentration gradient on the right. Reprinted with permission from C. Tanford, *Physical Chemistry of Macromolecules*, Wiley, New York, 1961, p. 354.

Here again, the derivative on the left is carried out with x held constant and that on the right with t held constant.

For the boundary spreading experiments in which diffusion is studied, the concentration of the macromolecule as it spreads can be detected by various optical techniques. A schematic depiction of the boundary spreading is shown in Figure 18-2. An analysis of the boundary spreading as a function of time according to Fick's laws permits the determination of the diffusion constant.

On the surface, this appears to be an easy experimental measurement, but in actuality it is quite difficult. Ideally, an infinitely sharp boundary is set up at time zero, and boundary spreading is measured as a function of time in the absence of forces other than the concentration gradient. Diffusion is slow for macromolecules so that the experiments take a long time, typically days. During the experiment, the equipment must be mechanically stable to prevent distortion of the boundary, and the temperature must be very carefully controlled to prevent thermal gradients. In addition, the net charge of the macromolecule must be kept as close to neutrality as possible to minimize electrostatic gradients. In practice, most modern day studies of diffusion are done in an analytical centrifuge with a synthetic boundary cell in which the boundary is created by layering buffer on the top of the solution. At low centrifuge speeds, little movement of the molecules occurs due to centrifugal force. Some typical diffusion constants for macromolecules are given in Table 18-2.

Complex hydrodynamic calculations show that the diffusion constant, D , can be written as

$$D = \frac{k_B T}{f} \quad (18-6)$$

where k_B is Boltzmann's constant, T is the temperature, and f is the translational frictional coefficient. Thus, the frictional coefficient can be directly calculated from

TABLE 18-2. Selected Diffusion Constants in Water at 293 K

Sample	Molecular weight	$D \times 10^7$ (cm ² /s)	$\bar{V}_2$ (cm ³ /g)	f/f_{sphere}
Sucrose	342	45.9	—	—
Ribonuclease	13,683	11.9	0.728	1.14
Lysozyme	14,100	10.4	0.688	1.32
Bovine serum albumin	66,500	6.1	0.734	1.31
Hemoglobin	68,000	6.9	0.749	1.14
Tropomyosin	93,000	2.2	0.71	3.22
Myosin	493,000	1.10	0.728	3.65
Collagen	345,000	0.69	0.695	6.8
DNA	6,000,000	0.13	0.53	15
Tobacco mosaic virus	40×10^6	0.44	0.73	2.19

Data from Refs. 1 and 2.

the diffusion constant. If the molecule is assumed to be a sphere, f is given by Eq. (18-1), and Eq. 18-6 can be written as

$$D = \frac{k_B T}{6\pi\eta R_S} \quad (18-7)$$

The radius, R_S , can be related to the partial specific volume, $\bar{V}_2$, by the equation

$$R_S = \left[\frac{3M\bar{V}_2}{4\pi N_0} \right]^{1/3} \quad (18-8)$$

and

$$D = \left(\frac{k_B T}{6\pi\eta} \right) \left(\frac{4\pi N_0}{3M\bar{V}_2} \right)^{1/3} \quad (18-9)$$

The calculated diffusion constant for a given molecular weight will differ from the experimentally determined diffusion constant if the macromolecule is not spherical in shape. This is usually summarized by the ratio f/f_{sphere} , where f is the experimentally determined frictional coefficient and f_{sphere} is the calculated frictional coefficient for a sphere with a given molecular weight and partial specific volume (Eq. (18-8)). This ratio is included in Table 18-2. This ratio reflects both the shape and the hydration of the molecule. Typically, the effect of hydration is much less than the effect of shape. As expected, the ratio is near 1 for globular proteins (the first four entries for proteins), but deviates considerably for collagen, tropomyosin, and myosin (molecules known to be elongated) and DNA. These molecules are more accurately described as prolate ellipsoids rather than as spheres. An “apparent” radius can be calculated directly from Eqs. (18-1) and (18-7). If the molecule is not spherical, the apparent radius will be greater than the radius calculated from Eq. (18-8). Thus, the measurement of diffusion constants can give direct information about the hydrodynamic shape of a macromolecule.

18.4 CENTRIFUGATION

The most common laboratory technique that directly measures the hydrodynamic properties of molecules is centrifugation. Centrifugation involves generating a centrifugal field by rapidly spinning a sample in a rotor. The separation of a solid precipitate from a solution requires relatively slow spinning of the rotor. In order to cause macromolecules to move in a centrifugal field, very rapid spinning is required in a special instrument called an ultracentrifuge. The centrifugal force on a particle of mass m is $m\omega^2 r$, where ω is the angular velocity of the rotor in radians/s and r is the distance of the particle from the center of the rotor. The outward motion of the particle in the field will be partially opposed by the buoyancy of the molecule, that is, the mass of the solvent displaced by the macromolecule as it moves through the solution. The buoyant force is $\omega^2 r$ times the mass of the displaced solution, which is $m\bar{V}_2\rho$, where ρ is the density of the solvent. The net force on the particle, therefore, is the difference between these two opposing factors, namely, $\omega^2 rm(1 - \bar{V}_2\rho)$. The frictional force of the particles moving past each other with a relative velocity v is fv . When a particle moves in a centrifugal field in a closed cell, the net force is zero so that the centrifugal force on the particle is balanced by the frictional force and buoyancy:

$$fv = \omega^2 rm(1 - \bar{V}_2\rho) \quad (18-10)$$

and the velocity of the particle in the field is

$$v = \frac{\omega^2 rm(1 - \bar{V}_2\rho)}{f} \quad (18-11)$$

As expected, the more massive the particle, the more rapidly it sediments. Similarly, the larger the frictional coefficient, the slower the particle sediments. Also, the less dense the molecule and the more dense the solvent, the slower the particle sediments. In order to obtain a parameter independent of the speed of the centrifugation, it is customary to determine the *sedimentation coefficient*, s , the velocity per unit field:

$$s = \frac{v}{\omega^2 r} = \frac{m(1 - \bar{V}_2\rho)}{f} \quad (18-12)$$

The sedimentation coefficient has the dimensions of seconds and for typical macromolecules is in the range of 10^{-11} – 10^{-13} s. In order to eliminate the necessity of writing these exponentials, the sedimentation coefficient is usually reported in terms of the svedberg, S , with $S = 10^{-13}$ s. Theodor Svedberg (1884–1971) was the inventor of the ultracentrifuge and was awarded the Nobel Prize for this accomplishment.

18.5 VELOCITY SEDIMENTATION

Before proceeding further with discussion of the sedimentation coefficient, the instrumentation used for analytical ultracentrifugation is worth discussing. The rotor spins

at speeds greater than 50,000 revolutions per minute (rpm) [$\omega = 2\pi \text{ rpm}/60 \text{ radians/s}$], with forces greater than 600,000 times the force of gravity being produced. The rotor, therefore, must be mechanically stable and very carefully balanced. The cell holding the solution must be designed so that the optical properties of the solution can be observed during the centrifugation, typically through ultraviolet absorption. The design of the cell is shown in Figure 18-3. It consists of a centerpiece sandwiched by two quartz or sapphire windows. The centerpiece has two wedge-shaped sectors that are extensions of pie-shaped lines with the apex at the center of the rotor. With the windows in place, two cells are defined: one is filled with the solution of interest and the other with the solvent. The ultracentrifuge is equipped with an optical system that measures the absorbance of light at a range of distances, r , from the center of rotation. (Refractive index measurements also are possible.)

When the cell is spun, a boundary is formed between the fluid in each cell and an air bubble, creating a meniscus in each cell. Initially, the concentration of the solute is uniform throughout the fluid volume. As the centrifugation proceeds, a progressive movement of the macromolecule away from the meniscus occurs, and a moving boundary is formed between the solute and the trailing solvent. The net result is that the macromolecule moves to the rear of the cell. A typical plot of the concentration of the macromolecule versus the distance from the center of the rotor is shown in Figure 18-4. Note that the sample becomes diluted as the centrifugation proceeds because the cell is wider at the bottom (radial dilution). The boundary stays sharp as long as significant diffusion does not occur during the duration of the experiment, typically a few hours. If significant broadening of the boundary occurs during the

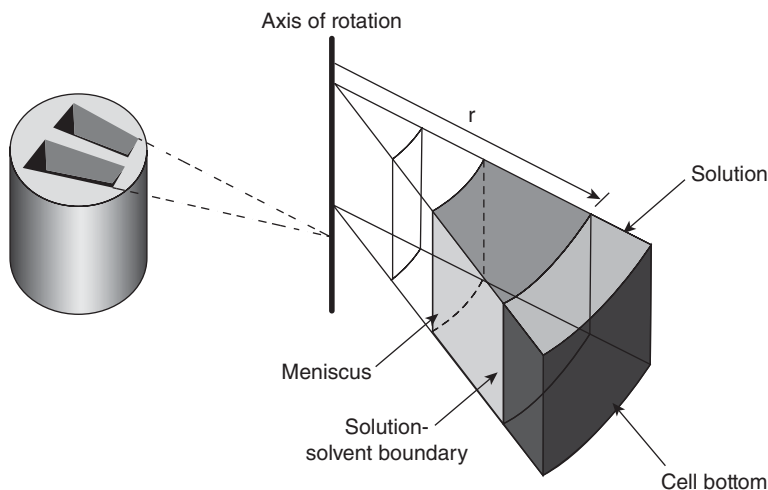


FIGURE 18-3. Schematic representation of the ultracentrifuge sample holder. The drawing on the left is a two-sector holder. One sector is used for the sample and the other for the buffer as an optical blank. Sample holders with more than two sectors are also used. The drawing on the right is a blowup of the sample sector only. The axis of rotation is shown to indicate that the sectors are designed as pie slices, with the axis of rotation at the center of the pie.

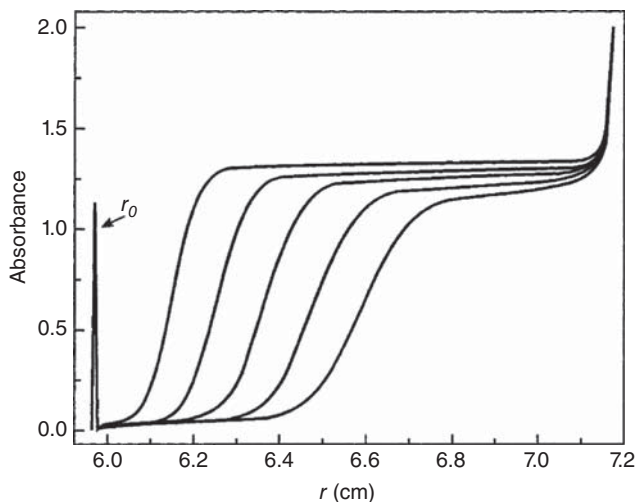


FIGURE 18-4. Boundary movement in a sedimentation velocity experiment. The absorbance is plotted versus r , the distance from the axis of rotation. The boundary moves to the bottom of the centrifuge cell as the centrifugation proceeds with a loss of sharpness due to diffusion. The meniscus at the top of the cell is at the position r_0 . The concentration decreases because the sector volume increases from the top of the cell to the bottom. The MoaC protein ($s = 5.21$ S) was centrifuged at 50,000 rpm and 20°C and the absorbance was scanned at 20 min intervals at 230 nm. Figure courtesy of Dr. Harvey Sage.

experiment, the diffusion constant can be calculated by analyzing the shape of the boundary as a function of time. Because diffusion is relatively slow for large macromolecules, diffusion constants are usually measured independently over a longer time period as previously discussed.

The sedimentation coefficient can be calculated from a plot of the position of the center of the boundary versus time. In practice, sedimentation coefficients are normally tabulated at a standard temperature of 20°C in water. If the sedimentation is measured at a different temperature, T , and solvent, the assumption is made that the only correction needed is to the frictional coefficient and specific volume of the macromolecule (Eq. (18-12)). If the frictional coefficient is assumed to depend only on the viscosity (Eq. (18-1)):

$$s_{20,w} = s_T \left(\frac{\eta_T}{\eta_{20,w}} \right) \frac{(1 - \bar{V}_2 \rho)_{20,w}}{(1 - \bar{V}_2 \rho)_T} \quad (18-13)$$

To make this conversion, the density and viscosity of the solvent used for the sedimentation experiment must be known, and the implicit assumption is that the shape and solvation of the macromolecule are the same in water and the solvent. In addition, the sedimentation coefficient is sometimes concentration dependent because of interactions between the macromolecules. In this case, the experimental sedimentation constant is extrapolated to zero concentration to obtain the value of $s_{20,w}$. From

TABLE 18-3. Selected Protein Sedimentation Coefficients in Water at 293 K

Sample	Molecular weight	s (S)	$\bar{V}_2$ (cm ³ /g)	f/f_{sphere}
Ribonuclease A	12,400	1.85	0.728	1.29
Lysozyme	14,100	1.91	0.688	1.22
Bovine serum albumin	66,500	4.31	0.734	1.33
Hemoglobin	68,000	4.31	0.749	1.28
Tropomyosin	93,000	2.6	0.71	2.65
Myosin	570,000	6.43	0.728	3.63
Tobacco mosaic virus	40×10^6	192	0.73	2.65

Data from Ref. 1.

Eq. (18-12), it is apparent that if the molecular weight of the species is known, the frictional coefficient can be calculated. As with diffusion, this provides information about the shape of the macromolecule. Sedimentation coefficients for some typical proteins are presented in Table 18-3, along with the ratio of the frictional coefficients, f/f_{sphere} , calculated from the sedimentation coefficients and diffusion constant.

The molecular weight of the macromolecule can be calculated from sedimentation velocity experiments if both the diffusion constant and the sedimentation coefficient have been measured. Equations (18-6) and (18-12) can be combined to eliminate the frictional coefficient. The result is

$$M = \frac{RTs}{D(1 - \bar{V}_2\rho)} \quad (18-14)$$

(Note that $m = M/N_0$ and $R = N_0k_B$.)

18.6 EQUILIBRIUM CENTRIFUGATION

The best method for determining molecular weights with ultracentrifugation is by equilibrium centrifugation. These measurements provide direct determination of the absolute molecular weight. As the macromolecule sediments in the centrifuge, a concentration gradient is created in the solution as the macromolecule accumulates at the back of the cell. This will cause the macromolecule to diffuse back in the direction where the concentration is depleted. Thus, diffusion causes the molecule to move to the top of the cell, and centrifugation causes it to move to the bottom of the cell. Eventually, these two opposing factors will balance, and a stable concentration distribution will be established within the cell. Typically, this takes many hours, even days, to occur. The final equilibrium distribution, however, has nothing to do with the shape of the macromolecule. Since the system is at equilibrium, all that matters is the energy of the macromolecule in the centrifugal field. From our previous considerations, it is apparent that the force acting on the macromolecule is $(M/N_0)(1 - \bar{V}_2\rho)\omega^2r$, where r is the distance from the center of rotation. The energy is the negative of the integral of

the force times dr . Thus, the energy per mole, E , at a given position in the centrifugal field is

$$E = -\frac{M(1 - \bar{V}_2\rho)\omega^2 r^2}{2} \quad (18-15)$$

The concentration distribution can be calculated from the Boltzmann equation that prescribes the probability distribution of the concentration as a function of energy, namely, for two concentrations c_1 and c_2 with energies E_1 and E_2 , respectively,

$$\frac{c_2}{c_1} = \exp \left[\frac{(E_1 - E_2)}{RT} \right] \quad (18-16)$$

Substituting Eq. (18-16) into Eq. (18-15), we obtain

$$\ln \left(\frac{c_2}{c_1} \right) = \frac{M(1 - \bar{V}_2\rho)\omega^2(r_2^2 - r_1^2)}{2RT} \quad (18-17)$$

This result predicts that a plot of $\ln c$ versus r^2 is a straight line and that the slope of this line permits the absolute molecular weight of the macromolecule to be determined. From Eq. (18-17), the molecular weight is

$$M = \frac{2RT}{(1 - \bar{V}_2\rho)\omega^2} (\text{slope of } \ln c \text{ versus } r^2) \quad (18-18)$$

A typical plot of $\ln c$ vs r^2 is shown in Figure 18-5 for bovine serum albumin.

Analytical centrifugation is one of the primary tools for determining the molecular weight of a macromolecule. It is particularly important if the protein contains multiple subunits or if a complex between nucleic acids and proteins occurs. Ultracentrifugation provides a means of determining the molecular weight of the complex formed. In addition, if multiple species are in equilibrium, for example, with concentrations of both complexed and uncomplexed species present, information can be obtained about the equilibria involved.

18.7 PREPARATIVE CENTRIFUGATION

Thus far, we have discussed ultracentrifugation as an analytical tool. However, it is also useful as a preparative method. This use is based on the obvious fact that macromolecules of different molecular weights centrifuge at different rates, as in velocity sedimentation. In principle, one might envisage simply centrifuging a mixture of macromolecules in a centrifuge tube and then punching a hole at the bottom of the centrifuge tube to elute the solution from the tube, with different molecular weight macromolecules eluting at different places in the elution pattern. This is the general idea of preparative or zonal centrifugation, but effective separation is not usually possible in a homogeneous solution because of gravitational instability. If a mixture of

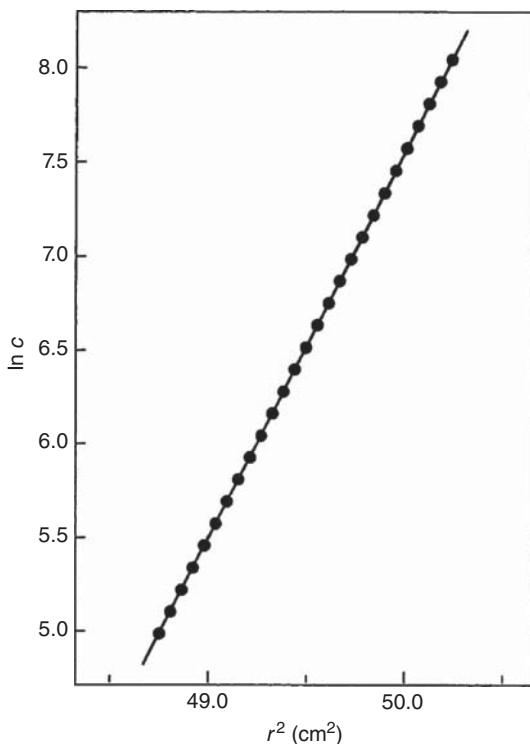


FIGURE 18-5. Sedimentation equilibrium data for bovine serum albumin in 0.1 M NaCl at 27,960 rpm, 20°C. The initial protein concentration was 650 $\mu\text{g}/\text{ml}$. The natural logarithm of the concentration is plotted versus the square of the distance from the axis of rotation (Eq. (18-17)). Reprinted from D. Freifelder, *Physical Biochemistry: Applications to Biochemistry and Molecular Biology*, 2nd edition, W.H Freeman, New York, 1982, p. 421. © 1976, 1982 by Freeman. Used with permission.

macromolecules is layered onto the top of a centrifuge tube, the density of the sample is higher in the layered band than in the solution below it, and the band would collapse. In zonal centrifugation, this problem is circumvented by creating a density gradient in the centrifuge tube before layering the sample onto the solution. This can be done with materials such as salt, sucrose, and glycerol. The density is highest at the bottom of the centrifuge tube. Particles will sediment through the gradient with a gradually decreasing velocity, each moving with an overall speed proportional to its s value. Thus, if the centrifugation is stopped before they reach the bottom of the tube, a separation of the macromolecules will occur, and the separate molecules can be isolated by collecting the contents of the centrifuge tube in fractions. If internal standards with known s values are included in the mixture of macromolecules, the values of the sedimentation coefficients for the individual macromolecules can be estimated.

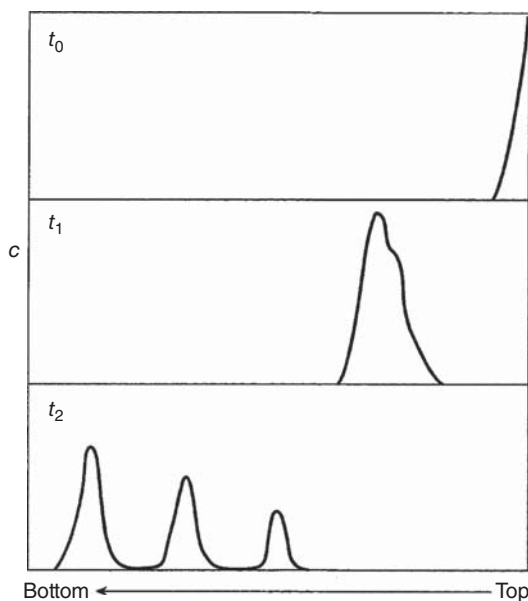


FIGURE 18-6. Schematic representation of zonal centrifugation of a mixture of three proteins. The proteins separate according to their s values as the centrifugation proceeds. In a real experiment, the area under the curves would remain constant.

The time course of zonal centrifugation for a mixture of three macromolecules is shown schematically in Figure 18-6. The separation of components depends on the molecular weights, the partial specific volumes of the macromolecules, and their shape.

18.8 DENSITY CENTRIFUGATION

Density centrifugation is related to zonal centrifugation and is also an effective preparative method. The best known example of this is the separation of nucleic acids by CsCl gradient centrifugation. Rather than creating a preformed gradient in a centrifuge tube, a CsCl solution is centrifuged until a CsCl gradient is formed. This is because small molecules will tend to distribute in the tube just as large molecules, and this creates an effective gradient if the solution is reasonably dense (a saturated solution of CsCl has a density of 1.9 g/cm^3). CsCl is commonly used because the solution density increases considerably as the concentration of CsCl is increased. However, other materials can work equally well, as long as a broad range of densities can be achieved. If a macromolecule is included in the solution, it will tend to form a tight band at a solvent density the same as its own, thus providing an effective method of separation because of the difference in buoyant density of the macromolecules. This has proved especially effective for the separation of different DNA molecules.

The most famous experiment done with density centrifugation was the original demonstration that DNA replicates by making a new complementary strand for each original strand in the parent DNA. This was done by growing bacteria with DNA having ^{14}N at all positions in media containing ^{15}N . In the original double-stranded DNA, each strand contained ^{14}N . In the first generation, the double-stranded DNA had one strand with ^{14}N and one strand with ^{15}N . This could be easily seen with density gradient centrifugation as the ^{14}N : ^{15}N DNA moved further from the center of rotation than the ^{14}N : ^{14}N DNA because it is heavier. In the second generation, some ^{15}N : ^{15}N DNA forms which moves even further down the centrifuge tube. This result, of course, is one of the basic tenets of modern molecular biology.

18.9 VISCOSITY

Viscosity plays a central role in hydrodynamics, as seen, for example, in the definition of the frictional coefficient. However, it is rarely used to determine the molecular properties of macromolecules because it provides limited information, and better tools are available. The measurement of viscosity is relatively easy. With a simple viscometer (e.g., Ostwald viscometer), the time, t , required for a defined volume of liquid to flow through a capillary is measured, and the flow time for an unknown liquid is compared with that of a liquid whose viscosity is known. The relative viscosities are related through the equation

$$\frac{\eta_u}{\eta_k} = \frac{\rho_u t_k}{\rho_k t_u} \quad (18-19)$$

where ρ is the density and the subscripts k and u are the known and unknown liquids, respectively. The most common unit used for the viscosity coefficient is the poise [$\text{g}/(\text{cm s})$] or centipoise (10^{-2} poise). The SI units are $\text{kg}/(\text{m s}) = \text{Ps}$, where P is the pascal.

The viscosity of a solution of macromolecules is strongly dependent on the concentration of the macromolecule. For a given concentration, the specific viscosity, η_{sp} , is defined as

$$\eta_{\text{sp}} = \frac{(\eta' - \eta)}{\eta} \quad (18-20)$$

where η' is the viscosity of the solution and η is the viscosity of the solvent. Finally, to obtain a viscosity that is intrinsic to the macromolecule and independent of the concentration of the macromolecule, c , the *intrinsic viscosity*, $[\eta]$, is defined as the value of η_{sp}/c extrapolated to zero concentration. Note that the intrinsic viscosity has the units of reciprocal concentration as the specific viscosity is dimensionless. This is usually taken as cm^3/g . The intrinsic viscosities of a few selected proteins are given in Table 18-4, along with the frictional coefficient ratio f/f_{sphere} determined by diffusion constant measurements.

TABLE 18-4. Intrinsic Viscosities of Proteins in Aqueous Salt (20–25°C)

Sample	Molecular weight	$[\eta]$ (cm ³ /g)	f/f_{sphere}^a
Ribonuclease	13,683	3.30	1.14
Bovine serum albumin	65,000	3.7	1.31
Hemoglobin	68,000	3.6	1.14
Tropomyosin	93,000	52	3.22
Myosin	493,000	217	3.65
Collagen	345,000	1150	6.8

Data from Ref. 2.

^aFrom diffusion constant measurements (Table 18-2).

The intrinsic viscosity can be related to the shape of the macromolecule through the relationship

$$[\eta] = \nu(\bar{V}_2 + \delta\bar{V}_1) \quad (18-21)$$

In this equation, ν is a dimensionless factor that is dependent only on the shape of the macromolecule. For a sphere, ν is 2.5, and it is substantially larger for nonspherical shapes. Values of ν are included in Table 18-1. From Eq. (18-21), it can be seen that the intrinsic viscosity is a measure of the shape and/or volume of a macromolecule. Thus, measurements of the intrinsic viscosity can be used to monitor processes such as the assembly of multimolecular species or large shape changes such as protein denaturation.

Viscosity measurements, for example, can be used to show that the activity of DNA polymerase increases the length of DNA. When the first polymerases were isolated, the activity of the enzyme was demonstrated by the incorporation of radioactive nucleotides into DNA. However, this could be due to synthesis of new DNA or due to exchange reactions with existing DNA. The viscosity of DNA as it grows longer increases significantly so that viscosity measurements demonstrated that new DNA was being made.

18.10 ELECTROPHORESIS

Biological macromolecules are usually charged so that if an electric field is applied across a solution, the molecules will move in the field. Positively charged molecules will move to the cathode (negative electrode) and negatively charged molecules to the anode (positive electrode). The fundamental relationship for the movement of a molecule in a nonconducting solvent is

$$u = \frac{Ze\mathbf{E}}{f} \quad (18-22)$$

where u is the velocity of the molecule, Z is the ionic charge number of the molecule, e is the charge of an electron, $\mathbf{E}$ is the electric field, and f is the frictional coefficient.

(If mks units are used, u is in m/s, Z is dimensionless, e is in coulombs, $\mathbf{E}$ is in volts/m, and f is in kg/s.) In principle, this relationship can be used to determine the charge of a molecule, as was done for determining the charge of an electron; however, biological macromolecules are in water solutions with other ions that influence the ionic atmosphere around the macromolecule. This makes it virtually impossible to obtain quantitative information about macromolecules with electrophoretic methods. However, electrophoresis is a valuable analytical tool, and some of its applications will be briefly described.

The first major application of electrophoresis was the separation of the protein components in blood through electrophoresis in aqueous solutions. This groundbreaking experiment demonstrated that blood contained many different proteins that could be separated and purified. Most current applications, however, utilize electrophoresis in gels. The gel is a polymer dispersed in aqueous solution that forms a three-dimensional network. Typical examples are agarose gels, a polysaccharide from agar in water, and polyacrylamide gels. Most of the gel is water, typically 90%, but the three-dimensional network of the polymer constricts the flow of molecules significantly. Furthermore, in some cases the macromolecules may be too large to enter the network. The extent to which flow is inhibited can be controlled by the degree of cross-linking of the gel. If a macromolecule is put into a gel and an electric field applied, the macromolecule will migrate in the gel. Exactly how far the macromolecule migrates is determined by the hydrodynamics of the macromolecule (size and shape) and its charge (Eq. (18-22)).

DNA fingerprinting is an example of the analytical power of gel electrophoresis. All humans have similar DNA, but the exact base sequences are very specific for individuals. Much of human DNA does not code for proteins (>90%) and consists of repeating sequences of bases. The number and nature of these sequences differ for each individual, and such a region is used for DNA fingerprinting. These sequences can be isolated, and restriction enzymes exist that cut double-stranded DNA at specific places determined by the nucleotides in the sequence. A large number of restriction enzymes are readily available, and they can be used to cleave the fingerprinting region of the DNA to obtain different fragmentation patterns. The fragments are then separated by electrophoresis on agarose gels. The DNA separates on the gel according to its size because every residue has the same charge due to the phosphate portion of the DNA. The pattern of bands obtained with only a few restriction enzymes is unique for each person (except identical twins) and therefore can be used to distinguish between people.

The actual detection of the gel band is accomplished by denaturing the DNA on the gel, either with heat or chemically, and making radioactive DNA oligonucleotides complementary to the individual gel bands. The gel bands are transferred ("blotted") to a nitrocellulose membrane, usually by simple capillary action. The membrane is then soaked with the radioactive nucleotides. After washing, the radioactivity of the blotted bands is measured. This is termed a Southern blot, after the person who invented this technique, Edward M. Southern. The use of this analysis in criminal cases is quite extensive. Not only can this method identify individuals with an extremely high probability, but it can also identify related people. The probability

of the analysis being correct depends on how many restriction maps are used—the odds become overwhelmingly large with relatively few restriction enzyme maps. Two other gel electrophoresis blotting methods are commonly used, one with radioactive RNA oligonucleotides and the other with antibodies. These are called Northern and Western blots, respectively—a scientific joke to those who know the origin of these terms.

Nucleic acids have one charge per nucleotide so that separation of oligonucleotides according to molecular weight is relatively easy with electrophoresis. The situation is quite different for proteins as the number of charges varies between individual proteins and with pH, salt concentration, etc. Furthermore, the shapes of proteins also vary, so they may have quite different frictional coefficients. Nevertheless, gel electrophoresis of native proteins may provide a convenient separation method, even though the migration of the protein in the gel cannot be quantitatively interpreted. Sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) gets around this problem by the addition of SDS to the protein. Proteins bind about the same amount of SDS per monomer. As SDS is highly charged, the charge per unit mass of protein is about the same. Furthermore, if disulfide bonds are broken by reduction with a reagent such as 2-mercaptoethanol, the protein is denatured by SDS into a similar shape for all proteins. If SDS-PAGE is carried out for reduced proteins reacted with SDS (typically about 10^{-3} M), the extent of movement of the proteins is determined entirely by their molecular weight. The protein is usually detected by staining the gel with a dye, for example, Coomassie blue, after electrophoresis. Under controlled conditions, the intensity of the stain can be related to the concentration in the gel.

The logarithm of the molecular weight is plotted versus the mobility of the protein for a variety of proteins and gels in Figure 18-7. The straight line relationship is a convenient method for determining the approximate molecular weight of a denatured protein. This straight line can be represented by the equation

$$\log M = A - Bx \quad (18-23)$$

where M is the molecular weight, x is the distance the protein has migrated in the gel, and A and B are constants determined empirically. In practice, a set of standard proteins is included in the gel along with the protein of unknown molecular weight in order to determine A and B . However, this is an empirical method and is not a substitute for determining the exact molecular weight by a method such as equilibrium centrifugation. If the protein itself is highly charged, has an unusual shape when denatured with SDS, or does not bind the “typical” amount of SDS, SDS-PAGE can give misleading results.

These two examples illustrate the importance of gel electrophoresis as an analytical technique in modern biology. These methods are dependent on the hydrodynamic properties of the macromolecules, even though these properties are used in an entirely empirical way. More extensive discussions of electrophoresis are available elsewhere (1,3).

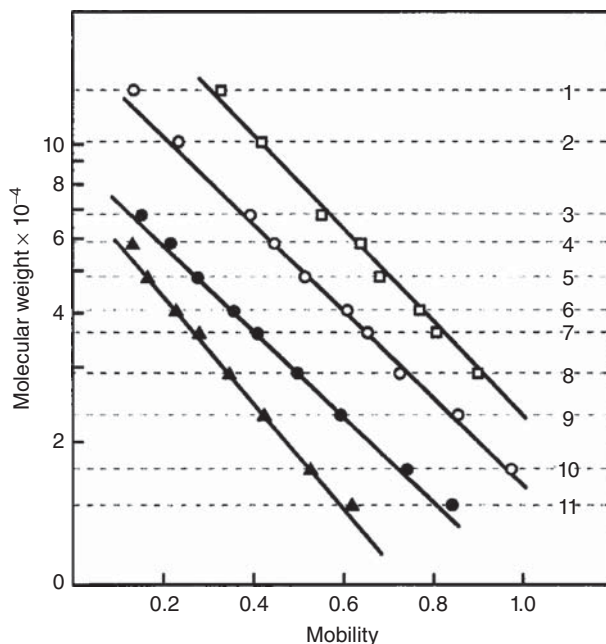


FIGURE 18-7. Electrophoretic mobilities of standard proteins in an SDS-PAGE experiment. The numbers refer to the standard proteins used, and the lines are different acrylamide concentrations. The acrylamide concentrations are 15% ($\blacktriangle$), 10% ($\bullet$), 7.5% ($\circ$), and 5% ($\square$). As expected, the higher the acrylamide concentration, the slower the migration. The standard proteins are β -galactosidase, phosphorylase A, serum albumin, catalase, fumarase, aldolase, glyceraldehyde phosphate dehydrogenase, carbonic anhydrase, trypsin, myoglobin, and lysozyme. Reprinted from K. Weber, J. R. Pringle, and M. Osborn, Measurement of molecular weights by electrophoresis on SDS-acrylamide gel, *Methods Enzymol.* **26**, 3 (1972). Copyright 1972, with permission from Elsevier.

18.11 PEPTIDE-INDUCED CONFORMATIONAL CHANGE OF A MAJOR HISTOCOMPATIBILITY COMPLEX PROTEIN

Major histocompatibility complex proteins (MHC) are cell surface proteins that are essential for the cell-mediated immune response. Class II MHC proteins bind peptides that are produced by proteolysis and present them to the surface of the T cell, thereby initiating an immune response (5). Complexes of these proteins with peptides have been extensively studied by a variety of techniques, including X-ray crystallography, thermodynamics, and kinetics. Here we briefly describe a hydrodynamic investigation that demonstrated a major conformational change of an isolated MHC protein induced by the binding of a peptide (6). In this work, a soluble portion of a specific MHC protein, DR1, was cloned and expressed in *E. coli*. DR1 consists of two different polypeptide chains that are tightly associated. The hydrodynamic properties of the protein and the protein-peptide complex were then measured by

TABLE 18-5. Hydrodynamic Data for DR1 and a DR1-Peptide Complex

	DR1	DR1-peptide
Gel filtration molecular weight	48,000	40,000
s (S)	2.67	3.23
D ($\times 10^{-7}$ cm ² /s), light scattering	6.17	7.40
D ($\times 10^{-7}$ cm ² /s), sedimentation velocity	5.73	6.74
D ($\times 10^{-7}$ cm ² /s), calculated		7.68
f/f_{sphere} , light scattering	1.3	1.1
R_s (Å), light scattering	35	29
R_s (Å), sedimentation velocity	37.4	31.8
R_s (Å), calculated		27.9

D is the translational diffusion constant, s is the sedimentation constant, R_s is Stokes' radius, and f/f_{sphere} is the ratio of frictional coefficients. The calculated values are based on the crystal structure of the DR1-peptide complex and the estimated hydrated volume of the DR1-peptide. Data from Ref. 8.

a variety of techniques: gel filtration, light scattering, and ultracentrifugation. The results obtained are summarized in Table 18-5 for DR1 and a DR1-peptide complex.

Gel filtration is an empirical method for determining molecular weight. Molecular sieve gels have pores in the individual beads that allow proteins to enter if they are small enough. The proteins partition between the inside and the outside of the beads. In general, the larger proteins partition more favorably on the outside of the beads, and the smaller proteins partition more favorably on the inside of the beads. A variety of gel materials are available with different pore sizes so that graded partitioning between the inside and the outside is available for a wide range of molecular weights. If proteins are passed through a gel filtration column in which all of the proteins can be partitioned inside the gel beads to some extent, they will elute from the column in the order of their molecular weight, with the highest molecular weight appearing first. If the column is calibrated with proteins of known molecular weights, the molecular weight of an unknown protein can be determined. Gel filtration, however, must be used with caution for determining molecular weights, as the shape of the macromolecule and its hydration are also important factors in determining the elution volume of a given protein. In fact, gel filtration separates proteins on the basis of their effective hydrodynamic radius rather than simply by molecular weight.

The diffusion constant obtained from the sedimentation coefficient was calculated by the use of Eq. (18-14), Stokes' radius with Eq. (18-7), and the frictional coefficient ratio as in Table 18-2. Table 18-5 includes a diffusion constant and a Stokes' radius calculated from the crystal structure of the DR1-peptide. This was done by calculating a radius based on the specific volume of the protein and the estimated water of hydration. The radius obtained from this procedure was multiplied by the frictional coefficient ratio for an oblate ellipsoid estimated from the crystal structure. In other words, $R_s = \left(\frac{3V_h}{4\pi} \right)^{1/3} (f/f_{\text{sphere}})$, where V_h is the hydrated volume.

The results obtained from the three methods are similar: binding of the two peptides causes an apparent decrease in the molecular weight and the Stokes' radius,

and an increase in the diffusion constant. This cannot be due to an actual change in molecular weight as the peptide does not significantly change the molecular weight of DR1. This was confirmed by sedimentation equilibrium experiments. The molecular masses were found to be 45,000 and 47,000 for DR1 and DR1-peptide, respectively. The most obvious way to explain these results is that the DR1-peptide complex has a more compact shape than the DR1 protein alone. Circular dichroism measurements suggested that structural changes occur when DR1 binds peptide. Thus, the evidence is quite convincing that a significant conformational change occurs when the peptide binds.

What is the biological significance of this conformational change? This is a matter of speculation, but the argument is made that this conformational change is important for presentation of the peptide to the T cell and perhaps for regulation of the immune response.

18.12 ULTRACENTRIFUGE ANALYSIS OF PROTEIN-DNA INTERACTIONS

Translin is a nucleic acid binding protein that appears to recognize sequences found at chromosomal break points. Chromosomal breakage and rejoining is associated with the development of malignancies (7). Consequently, the interaction of translin with DNA is of considerable biological significance. Translin is a protein of 228 amino acid residues and a molecular weight of 26,180 (calculated from the DNA sequence of its gene).

However, it exists and functions as a multisubunit protein. Ultracentrifugation analysis was carried out to determine its molecular weight and shape. In addition, the binding of oligonucleotides to translin was also investigated (8).

Velocity sedimentation analysis revealed a major species with an s value of 8.5 S. A broad distribution of a small amount of heavier material was present that is attributed to larger aggregates. The major species was attributed to an octamer with a frictional coefficient ratio of 1.35. In order to confirm the presence of an octamer and to examine more carefully the distribution of molecular species, equilibrium centrifugation experiments were also carried out. The results confirmed the presence of an octamer, along with a small amount of high molecular weight material. No evidence of lower molecular weight species was found, suggesting that the octamer is a very stable species. The shape of the octamer was derived from the frictional coefficient and electron microscopy. The octamer can be approximated by an oblate ellipsoid with an axial ratio of 7.5:1.

The interaction of translin with a 24-mer oligodeoxynucleotide was also studied. The oligonucleotide was labeled with the fluorescent molecule fluorescein. With this label, it was possible to observe both the free and unbound nucleotides by measuring the absorbance at 490 nm. Data were also collected at 280 and 260 nm: as discussed in Chapter 10, proteins have a maximum absorbance at 280 nm and nucleic acids have a maximum absorbance at 260 nm. However, at these two wavelengths, both the protein and the nucleic acid contribute to the absorbance. Equilibrium sedimentation

experiments were carried out for several concentrations of DNA and translin, and these data were simultaneously fit to a model in which the nucleic acid was assumed to form a one-to-one complex with the octameric translin. The equilibrium dissociation constant obtained was 84 nM. This is very tight binding, attesting to the specificity of this interaction.

The overall model that emerges from this work (and electron microscopy) is that translin forms an annular oblate ellipsoid structure of eight subunits that binds tightly to DNA at chromosomal break points.

As is apparent from the aforementioned discussion, hydrodynamic considerations form the basis of several important experimental methods used in modern biological research. Moreover, determination of the hydrodynamic properties of macromolecules provides important information about their structure and shape which in turn provides insight into their function.

REFERENCES

1. C. R. Cantor and P. R. Schimmel, *Biophysical Chemistry*, W. H. Freeman, San Francisco, CA, 1980.
2. C. Tanford, *Physical Chemistry of Macromolecules*, Wiley, New York, 1961.
3. D. Freifelder, *Physical Biochemistry: Applications to Biochemistry and Molecular Biology*, 2nd edition, Freeman, New York, 1982.
4. I. Tinoco, Jr., K. Sauer, & J. C. Wang, J. D. Puglisi, G. Harbison, and D. Rovnyak, *Physical Chemistry: Principles and Applications to the Biological Sciences*, 5th edition, Prentice Hall, Englewood Cliffs, NJ, 2013.
5. C. Watts, *Annu. Rev. Immunol.* 15, 821 (1997).
6. J. A. Zarutskie, A. K. Sato, M. M. Rushe, I. C. Chan, A. Lomakin, G. B. Benedek, and L. J. Stern, *Biochemistry* 38, 5878 (1999).
7. T. H. Rabbitts, *Nature* 372, 143 (1994).
8. S. P. Lee, E. Fuior, M. S. Lewis, and M. K. Han, *Biochemistry* 40, 14081 (2001).

PROBLEMS

- 18-1.** The value of s can usually be calculated by measuring the rate of change of the midpoint of the boundary, $x_{1/2}$, with respect to time (Fig 18-4). From Eq. (18-12), $s = v/(\omega^2 x) = (dx/dt)/(\omega^2 x) = (d \ln x/dt)/\omega^2$. Thus, the slope of a plot of $\ln x$ versus time and the value of ω allow s to be determined. The following data were obtained for a protein of molecular weight 74,000 with a specific volume of $0.737 \text{ cm}^3/\text{g}$ at 20°C . The density of the buffer is 1.010 g/cm^3 , the viscosity is 1.002 centipoise [10^{-2} g/(cm s)], and the centrifugation was carried out at 52,000 rpm.

Time (min)	$x_{1/2}$ (cm)
0	5.8591
30	5.9541
60	6.0763
90	6.2010
120	6.3282
150	6.4581
180	6.5905

- Calculate the sedimentation coefficient, s .
- Calculate the frictional coefficient.
- Calculate the diffusion constant.
- Calculate the frictional coefficient ratio, ff_{sphere} .

18-2. High molecular weight DNA can be modeled as a prolate ellipsoid of revolution. The frictional coefficient can be approximated as $f = 6\pi\eta a / [\ln(2a/b)]$, where η is the viscosity of the solvent, a is the long axis of the ellipsoid, and b is the short axis. DNA with a molecular weight of 1 million is about 5200 Å long and 22 Å in diameter.

- Calculate the diffusion constant and sedimentation coefficient in 0.1 M NaCl for this DNA at 20°C. The viscosity of the solvent at this temperature is 1.016 centipoise, its density is 1.0025 g/cm³, and the specific volume of DNA is 0.556 cm³/g.
- The calculated values are only in fair agreement with experimental values. Provide a possible explanation for this discrepancy.

18-3. Equilibrium centrifugation is carried out on a protein at pH 7.0 and pH 10.5 at 20°C. The partial specific volume of the protein is 0.749 cm³/g at this temperature and the density of water is 0.9982 g/cm³. The following results were obtained.

r^2 (cm ²)	c (μM) pH 7.0, 30,000 rpm	c (μM) pH 10.5, 40,000 rpm
49.00	0.431	0.333
49.10	0.611	0.388
49.20	0.865	0.453
49.30	1.22	0.529
49.40	1.72	0.616
49.50	2.44	0.720
49.60	3.46	0.840
49.70	4.89	0.980
49.80	6.91	1.14
49.90	9.78	1.33
50.00	14.2	1.55

- a. Calculate the molecular weight for the protein at the two pH values.
- b. Provide an explanation for the different molecular weights obtained at the two different pH values.
- 18-4.**
- a. Ribonuclease is a small molecular weight protein consisting of a single polypeptide chain. When it is thermally denatured, the relative viscosity of the solution increases. What is a likely explanation for this observation in terms of the molecular structures of the native and denatured states? What would you predict for the changes in the diffusion constant and sedimentation coefficients accompanying denaturation?
- b. In contrast, when double-stranded DNA is thermally denatured into single-stranded DNA, the viscosity of the solution decreases. Explain this observation in terms of the molecular structures of single- and double-stranded DNA. If the melting temperature of the DNA is 30°C, sketch a plot of the intrinsic viscosity versus temperature. In addition, what would you predict for the changes in diffusion and sedimentation coefficients for the DNA. (Assume that the degree of polymerization of the DNA is unchanged. DNA melting and the melting temperature are discussed in Chapters 3 and 6.)
- 18-5.** The ribosome from *E. coli* can be isolated as two particles, the 30S and 50S subunits. (The nomenclature for ribosomes uses the sedimentation coefficients determined in early experiments.) Both particles contain proteins and RNA. The 30S subunit contains 16S RNA and proteins. A question that was asked was whether the 16S RNA has the same configuration when it is isolated as when it is in the 30S subunit. One approach to this question is to determine the effective hydrodynamic radii of the isolated 16S RNA and the 30S subunit. The following data were obtained at 20°C [M. F. Tam, J. A. Dodd, and W. E. Hill, *J. Biol. Chem.* **256**, 6430 (1981)].

	$D (\times 10^7 \text{ cm}^2/\text{s})$	$s (\text{S})$	$\overline{V}_2 (\text{cm}^3/\text{g})$
16S RNA	1.721	21.0	0.541
30S subunit	1.97	31.8	0.590

- a. Calculate the molecular weights of 16S RNA and the 30S subunit. The density of water under these conditions is 0.9982 g/cm³.
- b. Calculate the effective Stokes' radii of the 16S RNA and the 30S subunit from the diffusion constant and sedimentation coefficient. The viscosity of water at this temperature is 1.005 centipoise. What do you conclude about the conformation of the isolated 16S RNA versus its conformation in the 30S subunit?

- c. Calculate the ratio f/f_{sphere} for 16S RNA and 30S subunit from the diffusion constant. What can you conclude about the shape of 16S RNA and 30S subunit. (The calculations in this problem are somewhat simplified relative to the original publication but the conclusions are the same.)

18-6. Aspartate transcarbamoylase is an enzyme that is important for the regulation of pyrimidines (Chapter 17). It has an unusual subunit structure that could have been determined by a combination of gel chromatography and SDS-PAGE. The hypothetical results of such experiments are summarized as follows. A gel filtration column was calibrated with the following proteins. The elution volumes can be assumed to be a linear function of the logarithm of the molecular weight.

Protein	Molecular weight	Elution volume (ml)
Cytochrome <i>c</i>	12,400	208
Myoglobin	17,800	194
Chymotrypsinogen	25,000	185
Bovine serum albumin	66,200	152
Lactate dehydrogenase	140,000	127
α -conarachin	295,000	102
Apo ferritin	475,000	90

- a. When aspartate transcarbamoylase was put through this column, it eluted with a volume of 101 ml. What is the molecular weight of aspartate transcarbamoylase?
- b. The enzyme was reacted with a reagent that modifies sulfhydryl groups and again passed through the gel filtration column. This time two peaks emerged with elution volumes of 138 and 174 ml. What are the molecular weights of these components?
- c. Finally, aspartate transcarbamoylase was subjected to SDS-PAGE. The following proteins were used to calibrate the gel.

Protein	Molecular weight	Gel migration (mm)
Lysozyme	14,400	88.0
β -Lactoglobulin	18,400	78.2
Carbonic anhydrase	29,000	70.0
Lactate dehydrogenase	35,000	55.2
Ovalbumin	45,000	46.0
Bovine serum albumin	66,200	30.2
β -galactosidase	116,000	10.0

Two bands were observed for aspartate transcarbamoylase with gel migrations of 58.0 and 82.0 mm. What are the molecular weights of these components?

On the basis of the aforementioned results, what can you say about the subunit structure of aspartate transcarbamoylase?

Mass Spectrometry

19.1 INTRODUCTION

The use of mass spectrometry in biology has become sufficiently prevalent that the basic concepts and their applications merit consideration in this text. The basic variable parameter in spectroscopy is wavelength, whereas in mass spectrometry, it is m/z , the ratio of the mass of a particle to its charge. In very simple terms, a mass spectrometry experiment can be divided into three main steps: ionization, mass analysis, and ion detection. Although modern mass spectrometry experiments are more complicated than suggested by this simple analysis, these three steps are always involved. Importantly, the only environment in which ions are stable for a sufficient time to analyze readily is a vacuum, and a good vacuum is an essential part of a mass spectrometer. This chapter constitutes an introduction to the field. More comprehensive treatments can be found in Refs. 1–4.

19.2 MASS ANALYSIS

The basic principles underlying mass spectrometry can be understood by considering the motion of a charged particle in electric and magnetic fields.

The kinetic energy of a charged particle in an electric field is given by the relationship

$$\frac{1}{2}mv^2 = zV \quad (19-1)$$

where m is the mass, v is the velocity, z is the charge, and V is the applied voltage. The trajectory of an ion in a magnetic field, H , is an arc of radius r , as shown in Figure 19-1. The conservation of angular momentum requires that

$$\frac{mv^2}{r} = zvH \quad (19-2)$$

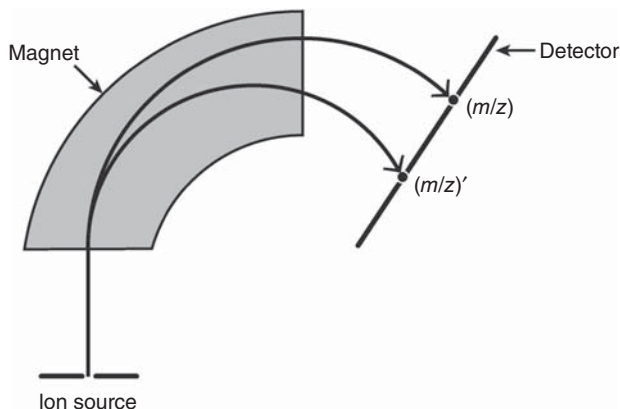


FIGURE 19-1. Schematic representation of the trajectory of ions with different m/z values in a magnetic field. In this illustration, $m/z > (m/z)'$.

Eliminating the velocity between these two equations gives

$$\frac{m}{z} = \frac{H^2 r^2}{2V} \quad (19-3)$$

This relationship shows that in constant magnetic and electric fields, the trajectory will be the same for all ions with the same ratio m/z . Thus, if the magnetic and electric fields are kept constant, the mass spectrum can be collected by determining the positions of the ions after they have passed through the magnetic field. Alternatively, the magnetic field can be scanned to determine the mass spectrum. Note that only the ratio m/z is determined, and multiply charged ions are commonly found.

The ultimate goal of the mass analysis is to determine m/z with the smallest possible error and to analyze a broad range of m/z . Unfortunately, these goals are often in conflict so that several different methods of mass analysis are used. The resolution of a mass spectrometer is defined as $m/\Delta m$. For example, a mass resolution of 1000 can distinguish between ions with m/z of 1000 and 1001.

The simplest type of mass spectrometer utilizes only a magnetic analyzer. As is evident from Eq. (19-2), the path of an ion in a magnetic field is different for every value of m/z . If Eq. (19-2) is solved for r , we find that

$$r = \left(\frac{1}{H} \right) \left(\frac{2mV}{z} \right)^{\frac{1}{2}} \quad (19-4)$$

The path for two ions with different m/z passing through a magnetic field is included in Figure 19-1. The resolution of the separation can be determined by taking the natural logarithm of Eq. (19-4) and differentiating:

$$\frac{\Delta r}{r} = \left(\frac{1}{2} \right) \left(\frac{\Delta m}{m} \right) + \left(\frac{1}{2} \right) \left(\frac{\Delta V}{V} \right) \quad (19-5)$$

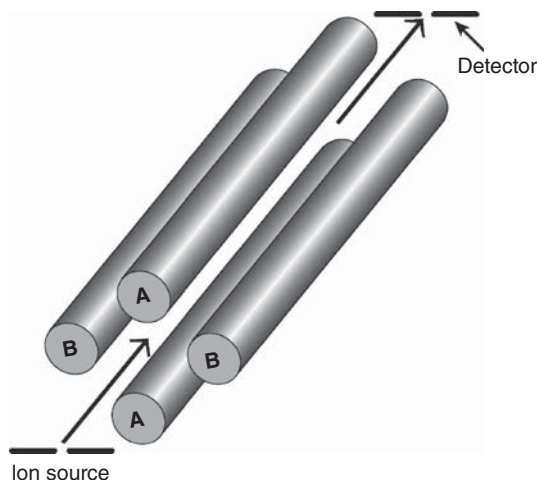


FIGURE 19-2. Schematic representation of a quadrupole mass analyzer. The ions pass through the middle of four parallel rods. The A rods are connected and have the same DC and superimposed radiofrequency voltages. The B rods are also connected and have a DC voltage opposite in sign to the A rods and the radiofrequency voltage phase shifted 180° relative to the A rods.

Thus, in order to sharply focus the ions, thereby decreasing the error in m , it is important for all ions to be homogeneous in energy, that is, the last term must be small. This can be partially accomplished by using a high voltage (V), typically 8000 V, but usually the magnetic focusing is coupled with an electrostatic analyzer in order to produce ions that are homogeneous in energy (small ΔV). This can be done by subjecting the ions to a constant voltage either prior or subsequent to the magnetic field. Instruments of this type, double-focusing magnetic sector mass spectrometers, have high resolution and a mass range up to about 15,000 Da. They are also very expensive.

The quadrupole mass spectrometer is a relatively low cost instrument in which the separation of ions is accomplished by utilizing electric fields only. The mass analyzer consists of four metal rods, as shown in Figure 19-2. The trajectory of the ions is between the four rods. The rods are electrically connected in pairs, A to A and B to B in the figure. A constant voltage of opposite sign is applied to the A and B rods. An oscillating voltage is superimposed on the constant voltage, with the phase differing by 180° between A and B. The quadrupole serves as a mass filter: only ions with a specified value of m/z can get through the filter. Other ions collide with the rods and do not reach the detector. To obtain a mass spectrum, the applied electric fields are varied, thus allowing ions with different values of m/z to be detected. The resolution of quadrupole instruments is not as good as double-focusing magnetic sector instruments, 10,000 versus 100,000, but it is significantly less expensive and can tolerate relatively high pressures. The latter feature is important if the ions are generated by electrospray, a frequently used technique that will be discussed later.

The quadrupole mass spectrometer has been coupled with an “ion trap.” With these instruments, the ions are held within the quadrupole and manipulated before proceeding to the detector. By holding the ions in the trap, both time and space dimensions are available. This improves the resolution and sensitivity. In simple terms, the ions are physically trapped between electrodes and subjected to both constant and oscillating electric fields such that ions of specific m/z precess within the trap. The magnitudes of the fields are increased, thereby causing ions of specific m/z to be ejected from the trap.

Modern instruments often couple linear ion traps or other ion sources with an orbitrap (5). In an orbitrap, the ions are trapped because of an electrostatic attraction to an inner electrode. When the voltage on the inner electrode is ramped up, the ions begin to orbit in spiral packets because the electrostatic interaction is balanced by a centrifugal force. Each packet has a unique m/z . These packets are detected on an outer electrode and analysed with Fourier transform methods.

Ion mobility spectrometry also has been coupled to mass spectrometry. In ion mobility spectrometers, ions drift down a tube in an electric field with an opposing flow of inert gas. Separation occurs on the basis of molecular size and shape as the ions move through the tube. A mass spectrometer is coupled to the end of this device to provide spectra of these preselected species.

The time-of-flight mass spectrometer selects ions by measuring the time of arrival of the ion at the detector. Equation 19-1 can be rearranged to give

$$v = \left(\frac{2V_z}{m} \right)^{\frac{1}{2}} \quad (19-6)$$

This equation predicts that the lighter the ion, the faster it will arrive at the detector. In order for this method to work, the ions must enter the flight tube at the same time. This is accomplished by generating ions in short bursts. The difference in the time of arrival of ions is not great, typically in the microsecond range. Thus, a complete spectrum can be determined in very short times. In some cases, reflectors are used to improve the sensitivity and resolution. This is done by slowing the ions with a series of electric field “lenses” until they essentially stop and then accelerating them in the opposite direction. The reflection increases the path length traveled, thus providing better separation of the time of flight, and the lenses focus the ions with a specific m/z by reducing the spread in kinetic energies for a given ion. The time-of-flight instrument has the advantage of essentially unlimited mass range, parallel detection of many ions, and high sensitivity.

Other methods exist for mass analysis but are not discussed in detail here. Most notably, Fourier transform methods using ion cyclotron resonance have been developed that permit very high-precision mass determinations. The ions are inserted into a small volume in the cyclotron, and a large magnetic field is applied so that the ions precess in circular orbits according to Eq. (19-2). The ions are constrained to the cell by an electric field applied to front and rear plates of the sample cell. If a pulsed electric field is applied at a frequency matching the precession frequency of the ions, energy is absorbed, analogous to a magnetic resonance experiment. The ions then transmit a radio frequency current at the detector plates that contains the frequency

TABLE 19-1. Characteristics of Mass Analyzers

Method	Mass range (Da)	Resolution
Magnetic sector	15,000	200,000
Quadrupole	4000	4000
Quadrupole ion trap	100,000	30,000
Time-of-flight	Unlimited	15,000
Fourier transform ion	$>10^6$	$>10^6$
Cyclotron resonance		

Adapted from reference 3.

components of each of the ions. This is converted to a free ion decay signal (analogous to the free induction decay in NMR) that can be Fourier transformed to the mass spectrum of all of the ions, thus permitting detection over a wide mass range.

A summary of the various mass analyzers, along with their approximate range of m/z and resolution, is given in Table 19-1.

19.3 TANDEM MASS SPECTROMETRY (MS/MS)

Tandem mass spectrometry couples two (or more) mass analyzers to obtain additional information about the sample in question. Three steps are typically involved in MS/MS analysis: mass selection, fragmentation, and mass analysis. The first step is the selection of a specific ion for further study. In the second step, the selected ion then undergoes fragmentation, usually through collisions with neutral gas atoms. The ion fragments are then analyzed by a second mass analyzer. Because modern ionization methods produce very little fragmentation, the ion selected is often similar to the ion of the parent compound. Selection of a specific ion and determination of its fragmentation pattern may be important for its molecular identification.

A variety of different MS/MS instruments are available. For example, a triple quadrupole instrument uses the first quadrupole for ion selection, the second for collision-induced dissociation, and the third for analysis of the fragments produced. Ion traps can also be incorporated at a relatively low cost. Similarly, multiple sector instruments are available that use a series of magnetic and electric fields for analysis. As might be expected, multiple sector instruments are quite expensive. Time-of-flight and Fourier transform cyclotron resonance instruments are also available. The use of MS/MS is particularly useful for biological systems to identify and characterize substances uniquely.

19.4 ION DETECTORS

The two most common methods of detecting ions after they have been sorted by the mass spectrometer are electron multipliers and photomultipliers. With electron multipliers, the ion strikes a dynode that emits secondary electrons. Typical dynode

surfaces are BeO, GaP, and CsSb. These secondary electrons are accelerated by a voltage and attracted to a second dynode that emits more electrons. This process continues through a series of dynodes, resulting in a cascade of electrons. The resulting current can be read with standard technology. Electron multipliers are very sensitive: a signal amplification factor of 10^6 can be readily obtained.

Photomultiplier detectors operate in a similar fashion except that the ion first strikes a phosphorous screen. The phosphorous screen releases photons, which are detected by a photomultiplier. The photomultiplier also has a series of dynodes and causes a cascade of electrons when the light strikes it. This type of detection is commonly called scintillation counting and is often used to measure radioactivity quantitatively. The amplification factor is similar to electron multipliers. A major advantage of photomultipliers is that they have significantly longer lifetimes than electron multipliers. The lifetime of electron multipliers is limited by contamination/damage of the surface that the ions strike. Nonetheless, electron multipliers are currently the most common devices used.

Detection of the ion signal is often done by a point detector so that only a single type of ion is detected, that is, a single m/z . However, array detectors are also available. Array detectors consist of a linear arrangement of detectors so that multiple ions can be detected simultaneously.

19.5 IONIZATION OF THE SAMPLE

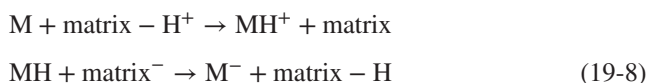
The first step in the analysis is to ionize the sample into the vacuum of the mass spectrometer. A variety of methods are used to ionize the sample, and we will deal with only a few of them. Currently, the two most widely used methods for macromolecules are MALDI (matrix assisted laser desorption/ionization) and ESI (electrospray ionization), but a few other methods that are commonly used for relatively small molecules will also be discussed.

Electron impact ionization vaporizes the sample into the mass spectrometer and then bombards the vapor with a beam of high-energy electrons, 50–100 eV. This is typically accomplished by thermal evaporation from a probe containing the sample. The high-energy electrons are produced from a filament and acceleration of the electrons through a large electric field. In order to produce ions, the electrons must have an energy greater than the ionization energy of the molecule, M , being studied. This process can be written as



The positive ions formed, M^+ , are referred to as odd-electron molecular ions or radical cations. Because the ionization energy for most molecules is only about 5 eV, the molecular ion produced usually has an excess of energy and fragments. The fragmentation pattern provides a “fingerprint” for each molecule. Libraries of fragmentation patterns are available for the identification of unknown compounds. The primary drawback of this method is that relatively stable and volatile compounds are required. In practice, this limits the molecular weight of samples to less than about 1000 Da.

The technique of fast atom bombardment (FAB) was developed in the 1980s and permits substances with molecular weights of 4000 or greater to be analyzed routinely. The principle of the method is to place a sample that is dissolved into a matrix on the tip of a metal probe. The sample is then bombarded with a stream of fast (high temperature) Ar or Xe atoms, and molecular ions are produced, predominately by the loss or gain of an H atom. (In some cases, a beam of Cs^+ is used.) The key to this method is the matrix, which is a viscous liquid and relatively inert. Typical examples are *m*-nitrobenzyl alcohol and glycerol. The sample is dissolved in a solvent that is miscible with the matrix. The matrix absorbs most of the high-energy atoms and produces a high-temperature, high-density gas within a small volume. Various matrix ions are produced, protonated, and deprotonated, and these react with the sample molecule, M, to produce ions:



Thus, both negative and positive sample ions are produced. The energy of the ions is considerably less than that produced by electron ionization, but some fragmentation occurs. A schematic representation of FAB is shown in Figure 19-3.

FAB has been largely replaced by MALDI, which is conceptually very similar (Fig. 19-3). The sample is embedded in a matrix, and an external source is used to convert the sample to ions. In the case of MALDI, the sample is embedded into a solid crystalline compound, and a laser, operating at a wavelength at which the matrix strongly absorbs, is used to desorb and ionize the sample. A laser with a high-energy output is needed: both UV and IR lasers have been used. The matrix is essentially a “solvent” for the sample, and the sample forms a microcrystal with the matrix. In practice, the sample is dissolved into a small amount of solvent (water or water plus

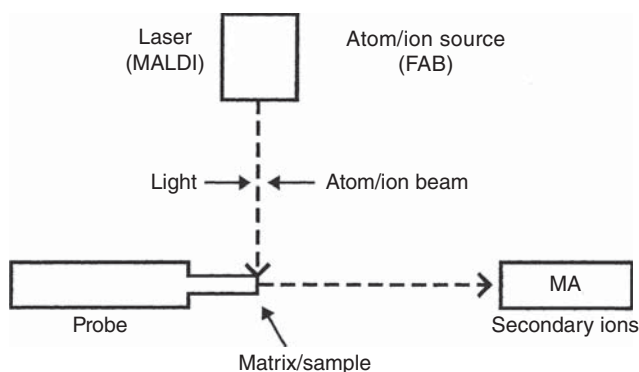


FIGURE 19-3. Schematic representation of FAB (right) and MALDI (left) methods of ionization. In both cases, the sample is embedded in a matrix. In the case of FAB, the ions are created by bombardment with an atom or ion beam, and in the case of MALDI by a pulsed laser. The secondary ions are then injected into the mass analyzer.

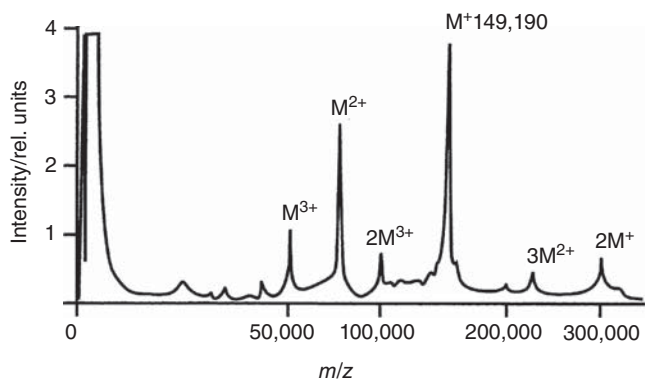


FIGURE 19-4. The MALDI mass spectrum of a monoclonal antibody. Reprinted from F. Hillenkamp and M. Karas, *Mass spectrometry of peptides and proteins by matrix assisted ultra-violet laser desorption/ionization*, *Methods Enzymol.* **193**, 280 (1990). © 1990, with permission from Elsevier.

organic solvents, typically), and the sample is then mixed with a concentrated solution of the matrix. Just a picomole of sample is required, and the ratio of sample to matrix is typically 1/1000. A variety of materials has been used for matrices, for example, cinnamic acid, succinic acid, and urea. The mixture is then evaporated and introduced into the mass spectrometer. Short laser pulses, typically 10–20 ns duration, with about 10^6 W/cm² power are used. The exact mechanism for producing ions is not well understood, but the general idea is that the matrix is ionized, and the ions produced ionize the sample molecules through a series of proton transfer reactions involving the matrix molecules and the analyte. As with FAB, both positive and negative ions are produced. This method can be used to study macromolecules with molecular weights in excess of 300,000 Da. Moreover, the sensitivity is excellent, in the pico- to femtomole range, and little fragmentation occurs. A typical mass spectrum of a protein obtained with MALDI is shown in Figure 19-4.

Finally, we consider ESI. With this method, a fine spray of highly charged droplets is created by spraying the sample onto the tip of a metal nozzle at approximately 4000 volts. This is done at atmospheric pressure. The solvent is removed through a series of Coulombic explosions. The droplets can also be heated to facilitate evaporation. As a droplet becomes smaller, the electric field density increases, until finally mutual repulsion between the ions causes the ions to leave the droplet. In essence, the electrostatic forces become greater than the surface tension. The ions are then directed into the mass analyzer with an electrostatic lens. An important and unique characteristic of this method is that highly charged ions are produced so that a wide range of m/z values are observed for a single analyte. For example, the mass spectrum of myoglobin obtained with ESI is shown in Figure 19-5. If two adjacent peaks are assumed to differ by only one charge and one proton, the molecular weight can be calculated. This method can be used for molecules of about 100,000 Da or less and has good sensitivity, in the pico to femtomole range.

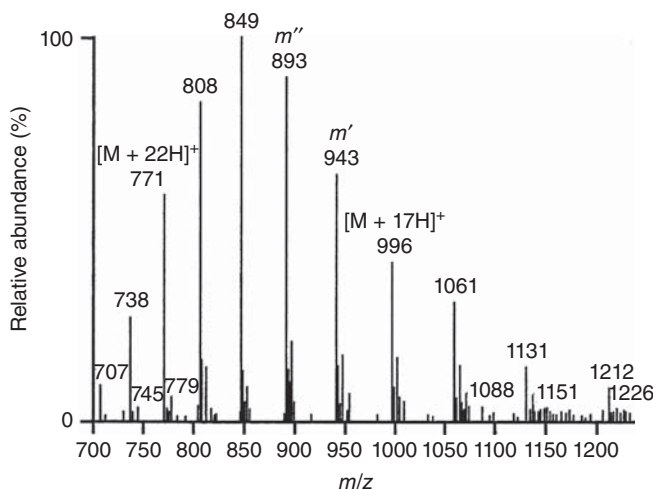


FIGURE 19-5. An electrospray ionization (ESI) mass spectrum of horse myoglobin. Each peak is characteristic of myoglobin with a different charge and number of protons. From C. Dass, *Principles and Practice of Biological Mass Spectrometry*, Wiley, New York, 2001, p. 43. Reprinted with permission of Wiley © 2001.

19.6 SAMPLE PREPARATION/ANALYSIS

Before discussing applications of mass spectrometry, some general comments should be made. Table 19-2 summarizes the most frequently used methodology for biomolecules. The amount of sample required depends on the specific methods used, but generally 5–50 μL of 10–100 μM solutions are adequate. Frequently, a mass spectrometer is coupled to purification procedures such as gas chromatography, high-performance liquid chromatography, or capillary electrophoresis. This provides a high throughput of samples and rapid analysis of unknown mixtures. In some cases, an in-depth study of a single substance is carried out. In this case, the purity is

TABLE 19-2. Ionization Techniques for Biomolecules

Compound	Ionization method	Ionization mechanism
Peptides	FAB, MALDI, and ESI	Protonation, deprotonation
Proteins	MALDI and ESI	Protonation
Carbohydrates	FAB, MALDI, and ESI	Protonation, deprotonation, cationization ^a
Oligonucleotides	MALDI and ESI	Protonation, deprotonation, cationization ^a
Small biomolecules	FAB, MALDI, and ESI	Protonation, deprotonation, cationization ^a , electron ejection

Adapted from G. Sluzdak, *Mass Spectrometry*, Academic Press, San Diego, CA, 1996.

^a Addition of a cation other than H^+ .

important and extraneous background material should be eliminated, notably salts, if possible.

In order to obtain accurate mass measurements with mass spectrometry, the instrument must be calibrated with standard compounds. This is sometimes done by running a series of calibration curves, but the best procedure is to include an internal mass standard with the sample. This will ensure that fluctuations in instrument response are adequately taken into account. Mass measurements can be made to within 1–500 parts per million. Thus, for example, the accuracy can be within a few tenths of a dalton for a macromolecule of molecular weight of 100,000.

19.7 PROTEINS AND PEPTIDES

With the advent of soft ionization methods, the determination of the molecular weight of proteins with mass spectrometry has become routine and has become the method of choice because of its accuracy and the relatively small amount of material required (femtomoles). The method works best if a pure protein is obtained by conventional methods such as chromatography. With a pure protein, mass spectrometry, usually with MALDI or ESI, can detect differences in a single amino acid quite readily. Thus, for example, mass spectrometry is a confirmatory tool for site-specific mutagenesis. Also if posttranslational modifications occur, such as phosphorylation or carbohydrate addition, mass spectrometry can identify and in some cases quantify the modification.

Proteomics is a rapidly developing field that systematically characterizes gene products. Proteomics is concerned with determining the identities of proteins, both known and unknown, and their functions. Tissues or fluids are subjected to protein purification protocols, usually by simple chromatography or gel electrophoresis, and the molecular masses and amino acid sequences of the proteins are then determined by mass spectrometry. This represents the first step in the characterization of gene products. Molecular weight by itself is rarely sufficient to identify a protein so that once proteins have been identified and their molecular weight determined, further studies are carried out. The next step is to take a specific protein, which could be a spot taken from two-dimensional gel electrophoresis of the crude starting material, and subject it to proteolysis with an enzyme such as trypsin. Before the use of mass spectrometry, the resultant peptides had to be separated, often by laborious procedures. However, with mass spectrometry, the trypsin digest can be analyzed directly and the molecular weights of the peptides determined. Since proteins usually have a unique digest, this procedure is often sufficient to determine what the starting protein is. Data banks of trypsin digests for hundreds of proteins and sophisticated software are available to help in the identification process. An example of the mass spectrum of a mixture of peptides is shown in Figure 19-6.

Although molecular weight characterization of peptide fragments after proteolysis is very useful for identifying proteins, the surest identification is to determine the amino acid sequence of the peptides. The amino acid sequence of the protein can then be compared with the vast database of known proteins. In addition to identifying

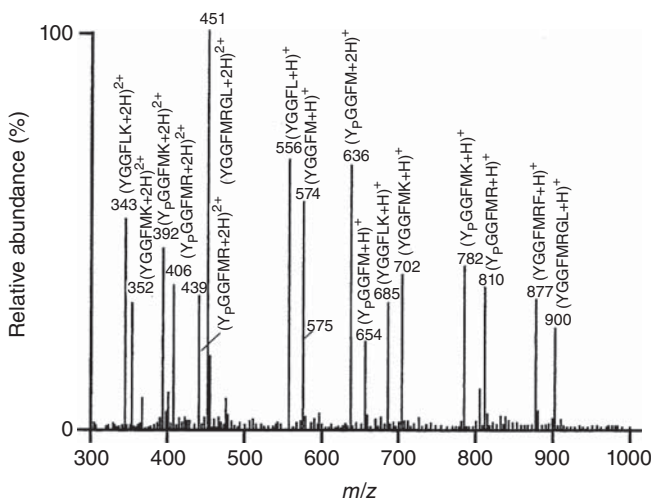


FIGURE 19-6. The positive ion ESI mass spectrum of a mixture of 10 peptides. The masses and amino acid sequences of the major peaks are indicated. Reproduced with permission from C. Dass in B. S. Larsen and C. N. McEwen (eds), *Mass Spectrometry of Biological Materials*, Marcel Dekker, New York, 1998, pp. 247–280.

specific proteins, the databases can be searched for homologous proteins, that is, proteins with similar but not identical amino acid sequences.

Determining the sequence of a peptide with mass spectrometry is not as easy as making molecular weight measurements. Basically, the principle is to use fragmentation of the parent peptide as a unique identifier of the sequence. The most frequently used technology is tandem mass spectrometry. With this method, a single peptide can be selected by the initial ion separation, and this peptide can then be subjected to fragmentation and analysis by the second mass analyzer. The amino acid sequence of peptides with molecular weights of up to about 3000 can be readily determined with this technique. If the peptide is large, the amount of useful sequence information that can be generated in the mass spectrometer is limited because of the complexity and large number of fragments produced.

An effective strategy for amino acid sequencing is the use of peptide ladders. With this methodology, a peptide is treated with an exopeptidase (carboxypeptidase or aminopeptidase). These enzymes take off one amino acid at a time from the C- or N-terminus. The products of these digestions are then analyzed with mass spectrometry, and the reduction in molecular weight can be used to deduce the amino acid sequence. An example of ladder sequencing is shown in Figure 19-7, where a peptide from HIV protease was sequenced by degradation from the amino terminus. The letter by each mass spectrum peak indicates the amino acid that was found on the N-terminus as the peptide decreased in length (6).

The above sequence of events has not taken into account the fact that proteins often contain disulfide linkages. If this is the case, the protein must first be subjected

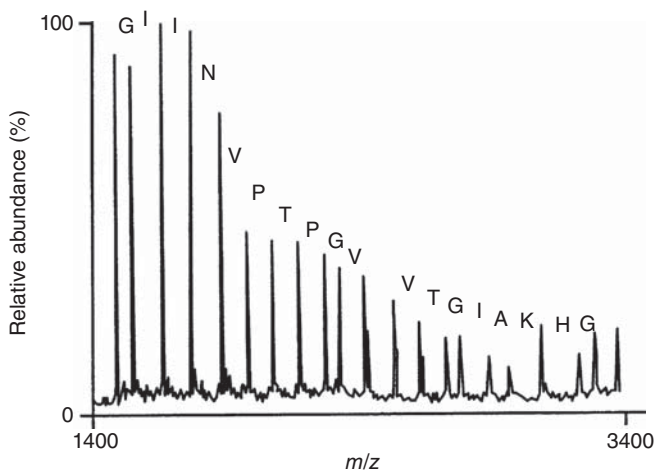


FIGURE 19-7. Mass spectrum of a mixture of peptides from an HIV protease following sequential removal of an amino acid from the N-terminus of the peptide. The letters indicate the amino acid that was removed from the N-terminus. Adapted with permission from B. T. Chait, R. Wang, R. C. Beavis, S. B. H. Kent, *Science* **262**, 89 (1993). © 1993 AAAS.

to a reduction process that converts the disulfides to cysteines, and the cysteines are then typically alkylated so that the disulfides will not reform. Because the molecular weights of the reduced/alkylated and oxidized enzymes are different, mass spectrometry can be used to determine exactly how many disulfide linkages are present.

Mass spectrometry can provide exact masses of proteins and other macromolecules. As the fields of genomics and proteomics expand, mass spectrometry undoubtedly will be a central analytical technique for the characterization of gene products. Both the methodology and database searching techniques are improving rapidly.

19.8 PROTEIN FOLDING

We have previously discussed the transformation of proteins between native and denatured states (Chapters 3, 11 and 14). Mass spectrometry provides a unique tool in such studies. Proteins have many hydrogens that can exchange with water hydrogens. This is a dynamic process and takes place continuously in aqueous solution. The amide hydrogens, that is, those associated with peptide bonds, are of particular interest because their rate of exchange with the solvent is dependent on structure and generally occurs over a time range that is readily accessible experimentally. All of the amide hydrogens are replaced by deuterium at approximately the same rate when a denatured protein is put into D_2O . However, for native proteins the rate depends on the environment of each amide hydrogen. For example, if the amide hydrogen is involved in a stable α -helix, the rate of exchange will be slower than that if it is freely

exposed to the solvent. Amide hydrogens directly exposed to the solvent exchange rapidly with the solvent. The remaining amide hydrogens can be divided roughly into two classes. One class exchanges due to the making and breaking of local structures: this is often called breathing motions of the protein. The second, slower, exchanging class of hydrogens is associated with the global structure of the protein. They usually do not exchange unless the native structure is denatured. Within each of these classes, subgroups can be found, especially the class associated with local structures (cf. Refs. 7 and 8 for reviews of this subject). For each subgroup, the exchange process can be described by the mechanism



Here, NH_{cl} and NH_{op} are the closed and open forms of the protonated enzyme: closed means hydrogen exchange cannot occur and open means hydrogen exchange can occur. ND is the deuterated amide(s), and the rate constants are for the opening, closing, and exchange of the unprotected amide reactions. If the intermediate state is assumed to be in a steady state, the rate constant for the overall exchange reaction is

$$k_{\text{ex}} = \frac{k_{\text{op}}k_{\text{int}}}{k_{\text{cl}} + k_{\text{int}}} \quad (19-10)$$

Two limiting cases exist. If the rate constant for closure is much faster than the rate constant for the replacement of hydrogen by deuterium ($k_{\text{cl}} \gg k_{\text{int}}$), $k_{\text{ex}} = k_{\text{op}}k_{\text{int}}/k_{\text{cl}}$. If $k_{\text{int}} \gg k_{\text{cl}}$, then $k_{\text{ex}} = k_{\text{op}}$. These are called the EX2 and EX1 limits. In the former case, because k_{ex} can be estimated from studies with model compounds, the equilibrium constant $k_{\text{op}}/k_{\text{cl}}$ can be calculated and is a direct measure of the thermodynamic stability of the structural element under consideration.

If deuterium is substituted for hydrogen, the molecular weight of the protein increases so that mass spectrometry is an excellent tool for studying hydrogen/deuterium exchange. NMR has also been used extensively because it can monitor specific individual amide protons in a single experiment. However, it cannot easily detect multiple populations of the same conformation. The combination of NMR and mass spectrometry provides a powerful experimental approach to the study of protein folding, but only a few selected examples studied with mass spectrometry are considered here.

In the EX2 limit, the open and closed forms of the protein are in equilibrium so that, as deuterium is substituted for hydrogen, only a single protein species of increasing molecular weight is observed during the time course of exchange. The starting protein species would have all amide hydrogens and the final protein species all deuteriums, with intermediate species having a specific number of hydrogens and deuteriums. This is shown schematically in Figure 19-8. In the EX1 limit, the opening reaction is rate determining, with rapid exchange following. Therefore, two species should be present during the time course of the exchange, one completely deuterated and the other completely hydrogenated at the amide position. This situation is also shown schematically in Figure 19-8. The shape of the mass spectrum peaks in

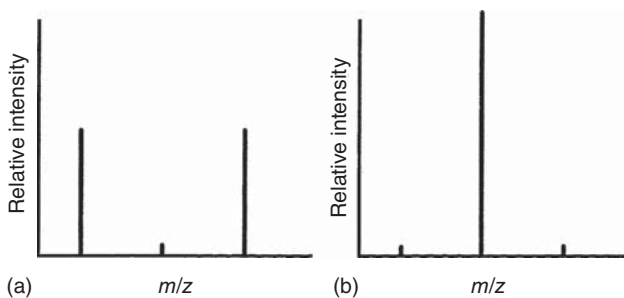


FIGURE 19-8. Schematic representation of mass spectra for the EX1 (a) and EX2 (b) exchange limits midway through the hydrogen/deuterium exchange reaction. The mass of the protein or peptide increases as the amide hydrogen is replaced by deuterium. In the EX1 limit, each protein or peptide molecule has either all hydrogens or all deuterons, whereas in the EX2 limit, each protein or peptide molecule contains both hydrogens and deuterons.

the EX1 limit can provide information about correlated exchange reactions and other reaction dynamics, but we will not discuss this more complex situation (8).

An extensive study of the incorporation of deuterium into rabbit muscle aldolase has been carried out (9). The protein was incubated in D_2O for a specified length of time. The exchange reaction was quenched by lowering the pD to 2.5 and the temperature to $0^\circ C$. Under these conditions, the rate of the exchange reaction (k_{ex}) becomes very slow. Denaturant was used to increase the initial rate of exchange—it promotes the formation of open forms. After quenching, the protein was proteolyzed with pepsin, and the peptides were subjected to mass spectral analysis by ESI. The time dependence of the exchange reaction was determined, and it was found that two different peptides incorporated deuterium at different rates. Both protonated and deuterated peptides were found, but intermediate states with both protons and deuterons were not found. This indicated that the EX1 limit is operative. Thus, the mass spectra showed that the EX1 mechanism of exchange was occurring and that the hydrogen/deuterium exchange for the two peptides took place at different rates, indicating differences in kinetic stability within the protein.

Mass spectrometry can also be used to study the exchange reaction in the EX2 limit which provides direct information about the thermodynamic stability of the protein. Mass spectrometry coupled with hydrogen exchange provides a means of rapidly scanning a library of proteins for relative stability (10). The stability of overexpressed proteins in crude cell lysates can also be assessed (11). In the latter case, the increase in mass of the overexpressed protein after exposure to D_2O for a fixed time was determined at different guanidine chloride concentrations. The reaction was quenched by addition to a MALDI matrix. The increase in mass for a number of variants of the protein λ repressor followed the typical sigmoidal behavior for a two-state denaturation, and the relative stability of seven mutants was determined. The results correlated well with those from circular dichroism measurements. The same method can be used to estimate binding constants as ligand binding generally stabilizes the native protein configuration. In addition, binding of a ligand or macromolecule may selectively alter

the exchange rate for specific regions of the target protein, thus providing information about the region of the protein involved in binding. The ability to rapidly screen many samples is a unique advantage of mass spectrometry for such studies.

19.9 OTHER BIOMOLECULES

Although the mass spectrometry literature has been dominated by the study of proteins in recent years, largely because of the genome project and proteomics, many other biomolecules have been studied. Posttranslational modification of proteins is a good example. After proteins are synthesized in the cell, they are often modified, for example, by glycosylation and phosphorylation.

The attachment of a carbohydrate to a protein is often important for its function and localization. Carbohydrates also play a significant role in the regulation of physiological processes. Determination of the structure of glycoproteins is challenging, and mass spectrometry can play an important role. If a protein is not homogeneous with respect to glycosylation, multiple molecular weights will be observed. MALDI is especially useful as it generally produces a single charged species and permits the multiple species to be easily sorted. ESI produces multiple ions for each species but has better mass resolution. The molecular mass of the attached carbohydrate can be determined by treating the protein with glycosidases to free the protein of the carbohydrate. If the extent of glycosylation is not great, the nature of the attached group can sometimes be inferred directly. If similar experiments are carried out in conjunction with peptide mapping (proteolysis and separation of peptides), the attachment sites of the carbohydrate can be ascertained. Ultimately, sequencing of the carbohydrate is required for complete characterization, and MS/MS methods are useful in this regard. The coupling of database searching and mass spectrometry is a promising approach for the characterization of carbohydrates. Although the characterization of carbohydrate structures is difficult and not yet well developed, its importance in biology is well recognized (1–3).

Finally, mention should be made of the application of mass spectrometry to characterizing oligonucleotides and lipids. In the case of lipids, it is a primary tool for structural characterization of the many diverse types of lipids that occur in nature. For nucleic acids, it is especially useful for determining modification of bases. We also have not dwelt on the characterization of metabolites; mass spectrometry is a tool of choice for small molecules because of the small amount of sample required to obtain a complete structure. The field of metabolomics subjects complex mixtures of metabolites obtained from physiological systems to mass spectrometry analysis to identify individual metabolites under different physiological conditions. A parallel field of lipidomics also exists that identifies individual lipids in complex mixtures.

REFERENCES

1. J. H. Gross and P. Roestorff, *Mass Spectrometry: A Testbook*, Springer Verlag, Berlin, Heidelberg, 2011

2. J. F. Watson and O. D. Sparkman: *Introduction to Mass Spectrometry: Instrumentation, Applications and Strategies for Data Interpretation*. John Wiley and Sons, New York, 2007.
3. C. Dass, *Principles and Practice of Biological Mass Spectrometry*, Wiley, New York, 2001.
4. G. Siuzdak, *The Expanding Role of Mass Spectrometry in Biotechnology*, MCC Press, San Diego, CA, 2003.
5. A. Makarov, *Anal. Chem.* 72, 1156 (2000).
6. B. T. Chait, R. Wang, R. C. Beavis and S. B. H. Kent, *Science* 262, 89 (1993).
7. T. C. Wales and J. R. Engen, *Mass Spectrom. Rev.* 25, 158 (2006)
8. D. M. Ferraro, N. D. Lazo and A. D. Robertson, *Biochemistry* 43, 587 (2004).
9. Y. Deng, Z. Zhang and D. L. Smith, *J. Am. Soc. Mass Spectrom.* 10, 675 (1999).
10. D. M. Rosenbaum, S. Roy and M. H. Hecht, *J. Am. Chem. Soc.* 121, 9509 (1999).
11. S. Ghaemmaghami, M. C. Fitzgerald and T. G. Oas, *Proc. Natl. Acad. Sci. USA.* 97, 8296 (2000).

PROBLEMS

- 19-1. The mass spectrum of the enzyme lysozyme was determined using ESI, and several positive ion peaks were observed. Peaks at $m/z = 1432$ and 1592 were found adjacent to each other. Calculate the molecular weight of lysozyme. Assume that the charges are due to protonation of the protein and that z differs by one unit for the two peaks.
- 19-2. A peptide was subjected to degradation from the N-terminus (Edman degradation). The resultant mixture was subjected to MALDI/MS analysis. The following ladder of m/z was observed: 977.2, 1064.3, 1151.4, 1222.6, 1378.9, and 1465.9. Assume that $z = 1$ in all cases and derive the sequence of the N-terminal region of the peptide. The identical peptide was enzymatically phosphorylated with protein kinase C. For this peptide, the ladder of m/z was as follows: 977.3, 1064.3, 1231.6, 1302.6, 1458.9, and 1545.9. Explain these results.
- 19-3. Derive an equation for the time of flight of an ion with a mass/charge ratio of m/z that travels a distance L in the mass spectrometer. Calculate the time of flight for a particle with $m/z = 200$ amu that was accelerated by 3000 volts over a distance of 30 cm. (*Hint*: Start with Eq. (19-6) in your derivation. Be careful of the units in making your calculation.)
- 19-4. Avidin is a glycoprotein found in egg white. It binds an important biomolecule, biotin, very tightly. Mass spectrometry of avidin using ESI displayed the usual multiple-peak spectrum due to multiple charges on a single avidin molecule. Two of the peaks had m/z values of 4002 and 4251. When the avidin was reacted with biotin prior to mass spectrometry, these peaks had m/z values of 4058 and 4312. Determine how many molecules of biotin are bound per molecule of avidin. The molecular weight of biotin is 244.

APPENDICES

Useful Constants and Conversion Factors

Avogadro's number, N_0	$6.0221 \times 10^{23} \text{ mol}^{-1}$
Gas constant, R	$8.3144 \times 10^7 \text{ erg K}^{-1} \text{ mol}^{-1}$
	$8.3144 \text{ J K}^{-1} \text{ mol}^{-1}$
	$1.9872 \text{ cal K}^{-1} \text{ mol}^{-1}$
	$0.0820571 \text{ atm K}^{-1} \text{ mol}^{-1}$
Boltzmann's constant, k_B	$1.3806 \times 10^{-23} \text{ J K}^{-1} \text{ molecule}^{-1}$
Planck's constant, h	$6.6262 \times 10^{-34} \text{ J/s}$
Speed of light, c	$2.9979 \times 10^8 \text{ ms}^{-1}$
Standard gravity, g	9.8066 ms^{-2}
Electronic charge, e	$1.6022 \times 10^{-19} \text{ C}$
Electron mass, m_e	$9.1094 \times 10^{-31} \text{ kg}$
Proton mass, m_p	$1.6276 \times 10^{-27} \text{ kg}$
Faraday constant, F	$9.6485 \times 10^4 \text{ C/mol}$
1 calorie = 4.184 Joule	
1 Joule = 10^7 erg = 1 Volt-Coulomb	
1 electron volt = 1.602×10^{-19} Joule	

Structures of the Common Amino Acids at Neutral pH

Aliphatic	$\begin{array}{c} \text{COO}^- \\ \\ ^+\text{H}_3\text{N}-\text{C}-\text{H} \\ \\ \text{CH}_3 \end{array}$ Alanine (Ala) (A)	$\begin{array}{c} \text{COO}^- \\ \\ ^+\text{H}_3\text{N}-\text{C}-\text{H} \\ \\ \text{CH} \\ / \quad \backslash \\ \text{H}_3\text{C} \quad \text{CH}_3 \end{array}$ Valine (Val) (V)	$\begin{array}{c} \text{COO}^- \\ \\ ^+\text{H}_3\text{N}-\text{C}-\text{H} \\ \\ \text{H}-\text{CH}_2 \\ \\ \text{H}_3\text{C}-\text{CH}-\text{CH}_3 \end{array}$ Leucine (Leu) (L)	$\begin{array}{c} \text{COO}^- \\ \\ ^+\text{H}_3\text{N}-\text{C}-\text{H} \\ \\ \text{H}_3\text{C}-\text{C}-\text{H} \\ \\ \text{CH}_2 \\ \\ \text{CH}_3 \end{array}$ Isoleucine (Ile) (I)
Nonpolar	$\begin{array}{c} \text{COO}^- \\ \\ ^+\text{H}_3\text{N}-\text{C}-\text{H} \\ \\ \text{H} \end{array}$ Glycine (Gly) (G)	$\begin{array}{c} \text{COO}^- \\ \\ ^+\text{H}_2\text{N}-\text{C}-\text{H} \\ \\ \text{H}_2\text{C} \quad \text{CH}_2 \\ \quad \quad \backslash / \\ \quad \quad \text{CH}_2 \end{array}$ Proline (Pro) (P)	$\begin{array}{c} \text{COO}^- \\ \\ ^+\text{H}_3\text{N}-\text{C}-\text{H} \\ \\ \text{CH}_2 \\ \\ \text{SH} \end{array}$ Cysteine (Cys) (C)	$\begin{array}{c} \text{COO}^- \\ \\ ^+\text{H}_3\text{N}-\text{C}-\text{H} \\ \\ \text{CH}_2 \\ \\ \text{CH}_2 \\ \\ \text{S} \\ \\ \text{CH}_3 \end{array}$ Methionine (Met) (M)
Aromatic	$\begin{array}{c} \text{COO}^- \\ \\ ^+\text{H}_3\text{N}-\text{C}-\text{H} \\ \\ \text{CH}_2 \\ \\ \text{C}=\text{CH} \\ / \quad \backslash \\ \text{N} \quad \text{NH} \\ \backslash \quad / \\ \text{C} \\ \\ \text{H} \end{array}$ Histidine (His) (H)	$\begin{array}{c} \text{COO}^- \\ \\ ^+\text{H}_3\text{N}-\text{C}-\text{H} \\ \\ \text{CH}_2 \\ \\ \text{C}_6\text{H}_5 \end{array}$ Phenylalanine (Phe) (F)	$\begin{array}{c} \text{COO}^- \\ \\ ^+\text{H}_3\text{N}-\text{C}-\text{H} \\ \\ \text{CH}_2 \\ \\ \text{C}_6\text{H}_4 \\ \\ \text{OH} \end{array}$ Tyrosine (Tyr) (Y)	$\begin{array}{c} \text{COO}^- \\ \\ ^+\text{H}_3\text{N}-\text{C}-\text{H} \\ \\ \text{CH}_2 \\ \\ \text{C} \\ / \quad \backslash \\ \text{C}_8\text{H}_6 \quad \text{NH} \end{array}$ Tryptophan (Trp) (W)

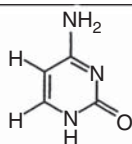
Physical Chemistry for the Biological Sciences, Second Edition.

Gordon G. Hammes and Sharon Hammes-Schiffer.

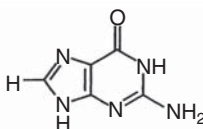
© 2015 John Wiley & Sons, Inc. Published 2015 by John Wiley & Sons, Inc.

Polar	$ \begin{array}{c} \text{COO}^- \\ \\ {}^+\text{H}_3\text{N}-\text{C}-\text{H} \\ \\ \text{CH}_2 \\ \\ \text{C} \\ // \quad \backslash \\ \text{O} \quad \text{NH}_2 \end{array} $	$ \begin{array}{c} \text{COO}^- \\ \\ {}^+\text{H}_3\text{N}-\text{C}-\text{H} \\ \\ \text{CH}_2 \\ \\ \text{CH}_2 \\ \\ \text{C} \\ // \quad \backslash \\ \text{O} \quad \text{NH}_2 \end{array} $	$ \begin{array}{c} \text{COO}^- \\ \\ {}^+\text{H}_3\text{N}-\text{C}-\text{H} \\ \\ \text{H}-\text{C}-\text{OH} \\ \\ \text{H} \end{array} $	$ \begin{array}{c} \text{COO}^- \\ \\ {}^+\text{H}_3\text{N}-\text{C}-\text{H} \\ \\ \text{H}-\text{C}-\text{OH} \\ \\ \text{CH}_3 \end{array} $
	Asparagine (Asn) (N)	Glutamine (Gln) (Q)	Serine (Ser) (S)	Threonine (Thr) (T)
Charged	$ \begin{array}{c} \text{COO}^- \\ \\ {}^+\text{H}_3\text{N}-\text{C}-\text{H} \\ \\ \text{CH}_2 \\ \\ \text{CH}_2 \\ \\ \text{CH}_2 \\ \\ \text{CH}_2 \\ \\ \text{NH}_3^+ \end{array} $	$ \begin{array}{c} \text{COO}^- \\ \\ {}^+\text{H}_3\text{N}-\text{C}-\text{H} \\ \\ \text{CH}_2 \\ \\ \text{CH}_2 \\ \\ \text{CH}_2 \\ \\ \text{N}-\text{H} \\ \\ \text{C}=\text{NH}_2^+ \\ \\ \text{NH}_2 \end{array} $	$ \begin{array}{c} \text{COO}^- \\ \\ {}^+\text{H}_3\text{N}-\text{C}-\text{H} \\ \\ \text{CH}_2 \\ \\ \text{C} \\ // \quad \backslash \\ \text{O} \quad \text{O}^- \end{array} $	$ \begin{array}{c} \text{COO}^- \\ \\ {}^+\text{H}_3\text{N}-\text{C}-\text{H} \\ \\ \text{CH}_2 \\ \\ \text{CH}_2 \\ \\ \text{C} \\ // \quad \backslash \\ \text{O} \quad \text{O}^- \end{array} $
	Lysine (Lys) (K)	Arginine (Arg) (R)	Aspartate (Asp) (D)	Glutamate (Glu) (E)

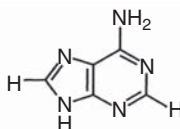
Common Nucleic Acid Components



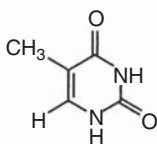
Cytosine
(C)



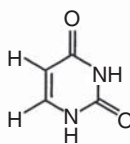
Guanine
(G)



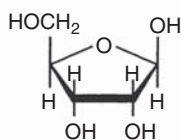
Adenine
(A)



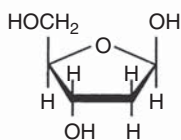
Thymine
(T)



Uracil
(U)



β -D-Ribofuranose



β -D-2-Deoxyribofuranose

Standard Gibbs Energies and Enthalpies of Formation at 298 K, 1 atm, pH 7, and 0.25 M Ionic Strength

Substance	ΔG° (kJ/mol)	ΔH° (kJ/mol)
ATP	−2097.89	−2995.59
ADP	−1230.12	−2005.92
AMP	−360.29	−1016.88
Adenosine	529.96	−5.34
P _i	−1059.49	−1299.39
Glucose-6-phosphate	−1318.92	−2279.30
Glucose	−426.71	−1267.11
H ₂ O	−155.66	−286.65
NAD _{ox}	1059.11	−10.26
NAD _{red}	1120.09	−41.38
NADP _{ox}	1011.86	−6.57
NADP _{red}	1072.95	−33.28
Acetaldehyde	24.06	−213.97
Acetate	−247.82	−486.83
Alanine	−85.64	−557.67
Ammonia	82.94	−133.74
Ethanol	62.96	−290.76
Pyruvate	−350.78	−597.04
Formate	−311.04	−425.55
Sucrose	−667.85	−2208.90
Total CO ₂	−547.10	−692.88
2-Propanol	140.90	−334.11
Acetone	84.89	−224.17

Physical Chemistry for the Biological Sciences, Second Edition.

Gordon G. Hammes and Sharon Hammes-Schiffer.

© 2015 John Wiley & Sons, Inc. Published 2015 by John Wiley & Sons, Inc.

Substance	ΔG° (kJ/mol)	ΔH° (kJ/mol)
Glycerol	−171.35	−679.84
Lactose	−670.48	−2242.11
Maltose	−677.84	−2247.09
Succinate	−530.62	−908.68
Fumarate	−523.58	−776.57
Lactate	−313.70	−688.28
Glycine	−176.08	−525.05
Urea	−39.73	−319.29
Ribulose	−328.28	−1027.12
Fructose	−426.32	−1264.31
Ribose	−331.13	−1038.10
Ribose 5-phosphate	−1219.22	−2042.40
Aspartate	−452.10	−945.46
Glutamate	−372.16	−982.77
Glutamine	−120.36	−809.11
Citrate	−966.23	−1513.66
Isocitrate	−959.58	—
<i>cis</i> -Aconitate	−802.12	—
Malate	−682.83	—
2-Oxoglutarate	−633.59	—
Oxalosuccinate	−979.06	—
Oxaloacetate	−714.99	—
Glycerol 3-phosphate	−1077.14	—
Fructose 6-phosphate	−1315.74	—
Glucose 1-phosphate	−1311.89	—
CO ₂ (g)	−394.36	−393.51
O ₂ (g)	0	0
O ₂ (aq)	16.40	−11.70
H ₂ (g)	81.53	−0.82
H ₂ (aq)	99.13	−5.02

This table is based on the conventions that $\Delta G^\circ = \Delta H^\circ = 0$ for the species H⁺, adenosine, NAD[−], and NADP^{3−} at zero ionic strength.

Data are obtained from R. A. Alberty, *Arch. Biochem. Biophys.* **353**, 116 (1998).

Standard Gibbs Energy and Enthalpy Changes for Biochemical Reactions at 298 K, 1 atm, pH 7.0, pMg 3.0, and 0.25 M Ionic Strength

Reaction	ΔG° (kJ/mol)	ΔH° (kJ/mol)
$\text{ATP} + \text{H}_2\text{O} \rightleftharpoons \text{ADP} + \text{P}_i$	-32.48	-30.88
$\text{ADP} + \text{H}_2\text{O} \rightleftharpoons \text{AMP} + \text{P}_i$	-32.80	-28.86
$\text{AMP} + \text{H}_2\text{O} \rightleftharpoons \text{adenosine} + \text{P}_i$	-13.55	-1.22
$2 \text{ADP} \rightleftharpoons \text{ATP} + \text{AMP}$	-0.31	+2.02
$\text{G6P} + \text{H}_2\text{O} \rightleftharpoons \text{Glu} + \text{P}_i$	-11.61	-0.50
$\text{ATP} + \text{Glu} \rightleftharpoons \text{ADP} + \text{G6P}$	-20.87	-30.39

Data from R. A. Alberty, *Arch. Biochem. Biophys.* **353**, 116 (1998).

Introduction to Electrochemistry

A6-1 INTRODUCTION

Electrochemistry is an active field of modern research. It is also important in a practical sense for batteries, solar cells, and other energy conversion devices. The world of biology involves many redox reactions, as well as ion transport through membranes and electrical conduction in the nervous system. This appendix is intended to serve as an introduction to the basic concepts underlying electrochemistry and their linkages to thermodynamics.

A6-2 GALVANIC CELLS

Many of the fundamental ideas and definitions underlying electrochemistry can be illustrated by considering the galvanic cell shown in Figure A6-1. In this cell, metallic zinc is inserted into a solution of ZnSO_4 in one beaker and metallic copper is inserted into a solution of CuSO_4 in another beaker. The two solutions are connected electrically through a salt bridge such as Na_2SO_4 held in the U tube with fritted disks or a gel. When an electrical connection is established through the two electrodes, a voltage of approximately 1.1 volts is observed if the concentrations of salts in the two beakers are equal. This voltage or electrochemical potential is designated by a script E , ϵ . The voltage results from Zn metal being converted to Zn^{2+} at the Zn electrode (oxidation) and Cu^{2+} being converted to Cu metal at the other electrode (reduction). The electrode at which oxidation occurs is called the *anode*, and the electrode at which reduction occurs is called the *cathode*. Electrons flow from the anode to the cathode, and to preserve electroneutrality anions flow from the CuSO_4 solution to the ZnSO_4 solution. The two electrochemical reactions occurring are:



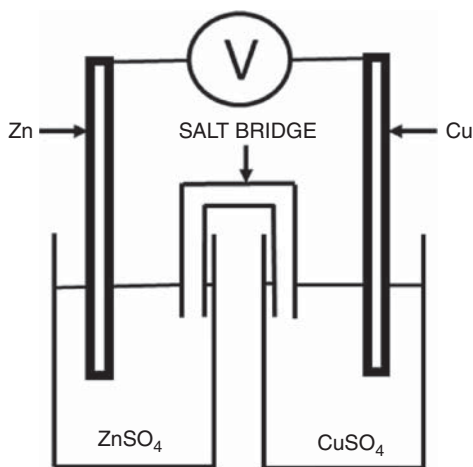
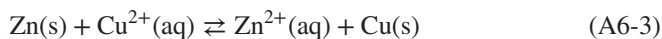


FIGURE 6-1. Galvanic cell with metallic Zn and Cu electrodes in solutions of ZnSO₄ and CuSO₄. The solutions are connected by a salt bridge containing Na₂SO₄. A voltage measuring device is inserted in the connection between the two electrodes. Galvanic cells today are more compact than this illustration and utilize modern materials.

The overall chemical reaction is



Why does the reaction proceed this way rather than the reverse? This is due to the intrinsic properties of the metal ions: Zn is more easily oxidized than Cu.

The convention for describing this cell is

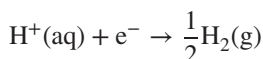


The single line designates a phase boundary, in this case metal/solution, and the double lines denote the salt bridge. The solution concentrations are shown within the parentheses.

The Gibbs energy associated with this cell is the electrical work done per mole, $-nF\epsilon$, where n is the number of electrons being transferred, and F is the Faraday constant. A word of caution: as the cell is described, it is not a reversible system. To make this a reversible system, an external voltage would have to be provided that just balances the voltage produced. The system is then at equilibrium with no net flow of electrons, and the usual thermodynamic concepts can be applied. Also this is a vastly simplified discussion of the galvanic cell. In reality potentials exist at all of the interfaces; however, we will not be concerned about these details.

A6-3 STANDARD ELECTROCHEMICAL POTENTIALS

Only differences in voltage can be measured experimentally, just as only differences in Gibbs energy can be measured. Moreover, the voltage observed is dependent on the ionic concentrations. To create an absolute scale of electrochemical potentials, two arbitrary assumptions are necessary. First, standard conditions must be established. The standard conditions are an activity of unity for the ions in the cell, a pressure of one bar, and constant temperature, usually taken as 25°C. The zero of the electrochemical potential is taken as the electrochemical potential for the reduction of hydrogen ion to hydrogen gas:



Making an electrode from hydrogen gas clearly is not feasible. Instead the electrode is made out of a material such as platinum and hydrogen gas is bubbled by it. Why platinum? First, it catalyzes the reduction reaction, and second, the electrochemical potential for the reduction of Pt is much less favorable (more negative) than for the reduction of $\text{H}_2(\text{g})$. This half cell can be represented as



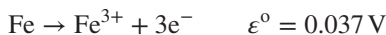
The standard conditions for this half cell are an activity of unity for $\text{H}^+(\text{aq})$ and a pressure of 1 bar for $\text{H}_2(\text{g})$. We have indicated that an activity of unity is the standard state. Obviously we cannot equate activity and concentration at such a high concentration so the assumption is made that activity and concentration are equal at low concentrations and the activity is then extrapolated to unity. We will not dwell on the procedure for doing this. The hydrogen electrode described above is no longer used. It has been replaced by a variety of glass electrodes that are sensitive to hydrogen ion concentration and can be used to measure pH.

Tables of standard electrochemical potentials, ϵ° , are available and can be utilized to calculate the standard electrochemical potentials of many different cells. For the reaction in Eq. (A6-1), ϵ° is 0.762 volts, and for the reaction in Eq. (A6-2), $\epsilon^\circ = 0.342$ volts. Thus the standard electrochemical potential for the galvanic cell discussed above is 1.104 volts.

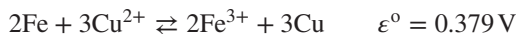
As a second example of an electrochemical cell, consider the following:



The two electrode reactions are



To write the overall reaction, the number of electrons must be balanced, so the first equation is multiplied by 2 and the second by 3 to give an equal number of electrons transferred for each half reaction. However, the standard voltages are not multiplied by these coefficients. The result is



The standard potential would be the same if the stoichiometric coefficients were divided or multiplied by integers: clearly the voltage cannot be dependent on how the reaction is written. On the other hand, the magnitude of the standard Gibbs energy change is dependent on how the reaction is written since $\Delta G^{\circ} = -nF\epsilon^{\circ}$. As the reaction is written, six electrons are transferred so $n = 6$.

A6-4 CONCENTRATION DEPENDENCE OF THE ELECTROCHEMICAL POTENTIAL

The concentration dependence of the electrochemical potential can be easily derived from previous thermodynamic considerations. The change in the partial molar Gibbs energy, ΔG , or chemical potential for a chemical reaction is related to the concentrations by the relationship (Eq. (2-35))

$$\Delta G = \Delta G^{\circ} + RT \ln Q \quad (\text{A6-4})$$

Here Q is the ratio of the product activities over the reactant activities with each activity raised to the power of its stoichiometric coefficient in the balanced chemical equation. From our discussion above, $\Delta G = -nF\epsilon$ and $\Delta G^{\circ} = -nF\epsilon^{\circ}$. If these relationships are substituted into Eq. (A6-4), the Nernst equation is obtained:

$$\epsilon = \epsilon^{\circ} - \frac{RT}{nF} \ln Q \quad (\text{A6-5})$$

This Nernst equation gives the dependence of the electrochemical cell on concentration. For the Zn/Cu electrochemical cell discussed above,

$$\epsilon = \epsilon^{\circ} - \frac{RT}{nF} \ln \left(\frac{[\text{Zn}^{2+}]}{[\text{Cu}^{2+}]} \right)$$

Here concentrations have been substituted for activities, and the activities of the metal electrodes have been set equal to 1. In practice, the substitution of concentrations for activities is valid up to about 10^{-3} m . Deviations from Nernst's law provide a method for determining activity coefficients. Nernst's law is frequently written in terms of the base 10 logarithm with $T = 298 \text{ K}$. In this case,

$$\epsilon = \epsilon^{\circ} - \left(\frac{0.0591}{n} \right) \log Q$$

Note that at equilibrium, $\varepsilon = 0$ and $\varepsilon^\circ = (RT/nF) \ln K$, where K is the equilibrium constant. Thus electrochemical cells can be used to determine equilibrium constants for redox reactions.

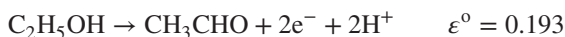
Finally we note that entropy and enthalpy changes for the electrochemical cell reaction can be determined from the temperature dependence of the electrochemical potential. Since $(\partial \Delta G / \partial T)_p = -\Delta S$, $\Delta S = nF (\partial \varepsilon / \partial T)_p$. The enthalpy of reaction can be obtained from the relationship, $\Delta H = \Delta G + T \Delta S$.

A6-5 BIOCHEMICAL REDOX REACTIONS

Thus far we have considered only redox reactions involving metals ions. However, the same basic concepts apply to all redox reactions. For example, alcohol dehydrogenase catalyzes the oxidation of alcohol to acetaldehyde by NAD^+ . The overall reaction is



The standard electrode potentials for the two half reactions at pH 7.0 are:



Biochemists use the convention that the activity of H^+ is unity at pH 7.0 so that H^+ does not explicitly appear in the equilibrium constants at pH 7.0. The standard potential for the alcohol dehydrogenase reaction is -0.131 V and $K = \exp(\varepsilon^\circ nF/RT) = [\text{CH}_3\text{CHO}][\text{NADH}]/([\text{C}_2\text{H}_5\text{OH}][\text{NAD}^+]) = 3.72 \times 10^{-5}$. Physiologically, the acetaldehyde is rapidly metabolized to acetate by aldehyde dehydrogenase. The acetate is ultimately broken down into H_2O and CO_2 to complete the pathway for the metabolic breakdown of alcohol,

More extensive discussions of electrochemistry and redox chemistry are available in the references given below.

REFERENCES

1. P. Atkins and J. de Paul, *Physical Chemistry* (9th edn), W. H. Freeman and Co., New York, NY, 2010.
2. R. Banerjee, D. Becker, M. Dickman, V. Gladyshev, and S. Ragsdale (eds), *Redox Biochemistry*, Wiley, New York, 2007.

INDEX

- ab initio* methods, 190–3
absorption spectra, 226–8
activation energy, 109
active transport, 52, 325
adenosine triphosphatase (ATPase), 14
 direct synthesis of, 49–51
adiabatic process, 10
allosterism, 408–11
 heterotropic interactions, 408
 homotropic interactions, 408
amino acids at neutral pH, structures of, 461–2
anode, 469
anomalous scattering, 210
anticooperativity, 400
anti-Stokes lines, 279
antisymmetry principle, 182
- Beer–Lambert law, 226–7, 230, 238
biochemical reactions, 43–4
biochemical redox reactions, 473
Bohr radius, 180
Boltzmann distribution, 338–43
Born interpretation, 158
Born–Oppenheimer approximation, 184–6, 357
Bragg equation, 208
Bravais lattices, 206
- calorimetry, 13–16
 simple batch calorimeter, 14
canonical distribution, 348
catalysis, 122
cathode, 469–70
centrifugation, 421
 density centrifugation, 427–8
 equilibrium centrifugation, 424–5
 preparative centrifugation, 425–7
chemical equilibria, 32–4, 87–9
chemical kinetics, 97–120
 first-order kinetics, 102
 principles of, 97–120
 pseudo-first-order kinetics, 103
 radioactive decay, 104–5
 rate constants, temperature dependence of, 108–12
 rate laws, determination of, 101–4
 reaction mechanisms, 105–8
 reaction rates, 99–101
 reaction rates near equilibrium, 114–16
 second-order kinetics, 103
 single molecule kinetics, 116–18
 thermodynamics and, 112–14
chemical potential, 80–83
chemical reactions
 mechanisms of, *See* reaction mechanisms
 membrane ion gradients establishment by, 51–2
chemical shifts, 293–6
chiral molecules, 253
 α -chymotrypsin, 126–33
 steady-state constants for, 127
circular dichroism (CD), 254, 256–7
 of nucleic acids, 259–60
 of proteins, 257–9
 protein secondary structure determined by, 282
circularly polarized light, 253
classical molecular dynamics, 358
closed system, 4
coherent scattering, 210

- constructive interference, 170
- conversion factors, 459
- cooperative process, 62
- cooperativity in ligand binding, 397–402
 - kinetic studies of, 406–8
 - models for cooperativity, 402–6
- COSY (Correlated Spectroscopy), 302
- Cotton effects, 256
- Cro repressor protein binding to DNA, 395–7
- crystal, scattering of X-rays by, 206–8

- de Broglie postulate, 157
- Debye–Hückel theory, 90–92
- degenerate states, 341
- degrees Celsius, 4
- density centrifugation, 427–8
- density functional theory (DFT), 193–4
- destructive interference, 170
- difference spectroscopy, 233–6
- diffusion, 418–21
- diffusion constant, 418
- dihydrofolate reductase (DHFR), 245–7, 286, 376–9
- DNA, 260–263
 - fingerprinting, 430
 - interaction with zinc finger proteins, 266–7
 - melting, 67–72, 142–8
 - small molecule binding to, 260–263
 - transcription regulation, 315–18
- docking, 364–5

- elastic scattering, 207
- electrochemistry, 469–73
- electron spin resonance (ESR), 306–10
 - usefulness in biological systems, 307
- electronic spectra, 225–51, *See also*
 - absorption spectra
 - difference spectroscopy, 233–6
 - dihydrofolate reductase (DHFR), 245–7
 - energy transfer application to biological systems, 243–5
 - fluorescence energy transfer, molecular ruler, 241–3
 - fluorescence, 236–40
 - nucleic acid spectra, 230–231
 - phosphorescence, 236–40
 - prosthetic groups, 231–3
 - RecBCD, helicase activity monitored by fluorescence, 240–241
 - ultraviolet spectra of proteins, 228–30
 - X-ray absorption spectroscopy, 236
- electronic structure, 177–202
 - atoms and molecules, 177–202
 - Born–Oppenheimer approximation, 184–6
 - density functional theory (DFT), 193–4
 - Hartree–Fock theory, 190–193
 - hydrogenic atoms, 177–80
 - many-electron atoms, 181–4
 - molecular orbital theory, 186–90
- electrophoresis, 429–32
- electrospray ionization (ESI), 446
- elementary reactions, 98–9
- elliptically polarized light, 253
- empirical valence bond (EVB) method, 374
- endopeptidase, 126
- energy, 3–21
 - definition, 9–10
- energy transfer application to biological systems, 243–5
- ensembles, 346–9
- enthalpy, 11–12
 - reaction enthalpies, 16–18
 - temperature dependence of, 18–19
- entropy, 23–42, *See also* second law of thermodynamics
 - calculation of, 26–8
 - molecular interpretation of, 29–30
 - statistical interpretation of, 349–50
- enzyme catalysis, 121–6, 219–22
 - Michaelis–Menten mechanism, 121–6
 - X-ray crystallography, 219–22
- enzyme reactions, simulations of, 376–9
- enzyme–substrate complexes, structure of, 286–7
- equilibrium approximation, 122
- equilibrium centrifugation, 424–5
- equilibrium dialysis experiment, 393
- ergodic hypothesis, 367
- Ewald summation method, 366
- exopeptidases, 126
- extended chemical potential, 39
- extensive properties, 4

- facilitated diffusion, 325
- fast ion bombardment (FAB), 447
- first law of thermodynamics, 9
- first-order kinetics, 102
 - pseudo-first-order kinetics, 103
- first-order saddle point, 196
- fluorescence, 225, 236–40
- fluorescence energy transfer, molecular
 - ruler, 241–3
- fluorescence polarization, 267–9
- Förster energy transfer, 241
- Fourier transform, 171
- Franck–Condon principle, 277
- free energy, *See* Gibbs energy and Helmholtz energy
- free induction decay (FID), 293
- frictional coefficient, 415–18

- galvanic cells, 469
- gauche conformations, 360
- Gibbs energy, 30–32, *See also* second law of thermodynamics
 - additions to, 39
 - calculations, 375–6
 - pressure dependence of, 35–6
 - temperature dependence of, 35–6
- global minimum, 196

- Hartree–Fock theory, 190–193
- heat, 5–6
- heat capacity, 5
- heat of formation, 12
- helix-coil transition, 350–353
 - initiation process, 350–351
 - propagation, 350
- α -helix protein structure, 57
- Helmholtz energy calculations, 375–6
- heme, 231–2
- Henry's law, 82
- Heteronuclear Single Quantum Correlation (HSQC), 303
- heterotropic interactions, 408
- HIV genome integration into host genome, 269–70
- homotropic interactions, 408
- host genome, HIV genome integration into, 269–70
- hybrid quantum/classical methods, 373–5

- hydrodynamics of macromolecules, 415–39
 - centrifugation, 421
 - density centrifugation, 427–8
 - diffusion, 418–21
 - electrophoresis, 429–32
 - equilibrium centrifugation, 424–5
 - frictional coefficient, 415–18
 - intrinsic viscosity, 428–9
 - peptide-induced conformational change of MHC protein, 432–4
 - preparative centrifugation, 425–7
 - ultracentrifuge analysis of protein–DNA interactions, 434–5
 - velocity sedimentation, 422–4
 - viscosity, 428–9
- hydrogen bond formation, 55
- hydrogenic atoms, 177–80
- hyperchromicity, 230
- hypochromicity, 230

- inelastic scattering, 207
- infrared (IR) spectroscopy, 278–9
 - protein secondary structure determined by, 282
- intensive properties, 4
- ion detectors, 445–6
- ionic solutions, 89–93
- ionization of sample, 446–9
- isobestic wavelength, 227
- isolated system, 4
- isomorphous replacement, 210
- isothermal process, 10

- α -ketoglutarate dehydrogenase, 270–272
- kinetic model of gases, 333–8
- kinetic studies of cooperative binding, 406–8
- kinetics applications to biological systems, 121–51, *See also* α -chymotrypsin; enzyme catalysis
 - DNA melting, 142–8
 - protein tyrosine phosphatase, 133–7
 - renaturation, 142–8
 - ribozymes, 137–42

Koshland, Nemethy, and Filmer (KNF)
model, 403–6

lactose permease, 325–8

Larmor frequency, 292

Lennard-Jones 12–6 potential, 363

ligand binding to macromolecules, 385–414

allosterism, 408–11

cooperativity in ligand binding, 397–402

Cro repressor protein binding to DNA,
395–7

experimental determination, 392–5

macroscopic equilibrium constants,
387–9

microscopic equilibrium constants, 387–9

models for cooperativity, 402–6

small molecules binding to multiple
identical binding sites, 385–7

statistical effects in, 389–92

linear polarized light, 253

Lineweaver-Burk equation, 125

local minima, 196

London attraction energy, 362–3

macromolecules, ligand binding to, 385–414

macromolecules, vibrations in, 277–88

enzyme–substrate complexes, structure
of, 286–7

infrared spectroscopy, 278–9

Raman spectroscopy, 279–81

vibrational spectroscopy, 281–3

macroscopic equilibrium constants, 387–9

macroscopic ionization constants, 388

magnetic resonance applications to biology,
315–29

DNA transcription regulation, 315–18

lactose permease, 325–8

proteasome structure and function, 328–9

protein folding, dynamics of, 320–322

protein–DNA interactions, 318–20

RNA folding, 322–5

magnetic resonance imaging (MRI), 306

major histocompatibility complex (MHC)

proteins, 432

peptide-induced conformational change
of, 432–4

many-electron atoms, 181–4

mass analysis, 441–5

characteristics of, 445

quadropole mass analyzer, 443

mass spectrometry, 441–55

ion detectors, 445–6

ionization of the sample, 446–9

mass analysis, 441–5

peptides, 450–452

protein folding, 452–5

proteins, 450–452

sample preparation/analysis, 449–50

tandem mass spectrometry (MS/MS),
445

mathematical tools, thermodynamics, 77–8

matrix assisted laser desorption/ionization
(MALDI), 446, 448

maximum velocity, 122

Maxwell distribution of speeds, 335–6

Maxwell relations, 78–80

melting process, DNA, 68

membrane ion gradients establishment by
chemical reactions, 51–2

metabolic cycles, 45–9

anaerobic glycolysis, 46–9

Metropolis Monte Carlo method, 373

Michaelis–Menten mechanism, 121–6

microscopic equilibrium constants, 387–9

microscopic ionization constants, 388

microscopic reversibility, 113

microstates, 23

minimum energy path, 196

molecular dynamics, 367–73

molecular interpretation of entropy, 29–30

molecular mechanical force field, 358

molecular orbital theory, 186–90

H₂, 188

He₂, 188

O₂ diatomic molecule, 190

molecular partition function, 343–6

molecular simulations, 357–81

enzyme reactions, simulations of, 376–9

Gibbs energy calculations, 375–6

Helmholtz energy calculations, 375–6

hybrid quantum/classical methods,
373–5

large-scale simulations, 365–7

molecular mechanics and docking, 364–5

potential energy surfaces, 358–64

Monod–Wyman–Changeux model, 403

Monte Carlo approach, 358, 373

multidimensional NMR, 300–306

- negative cooperativity, 400
- Nernst's law, 472
- neutron diffraction, 212
- nodes, 160
- NOE walk, 323
- normal modes of vibration, 163
- nuclear magnetic resonance (NMR), 289–313
 - chemical shifts, 293–6
 - magnetic resonance imaging (MRI), 306
 - multidimensional NMR, 300–306
 - NMR spectrometers, 292–3
 - relaxation times, 298–300
 - spin–spin splitting, 296–7
- Nuclear Overhauser Effect Spectroscopy (NOESY), 303–4
- nucleation, 145
- nuclei, magnetic properties of, 290
- nucleic acid components, 463
- nucleic acids, 259–60
 - circular dichroism of, 259–60
 - optical rotary dispersion of, 259–60
 - spectra, 230–231
- nucleic acid structures, 63–7, 213–15
 - DNA melting, 67–72
 - DNA/RNA chain, 63
 - RNA, 72
- open system, 4
- optical rotary dispersion (ORD), 254–6
 - experimental setup, 255
 - of nucleic acids, 259–60
 - of proteins, 257–9
- optical rotary dispersion, 253–76
- optical rotation, 254
- osmosis, 85
- osmotic pressure, 85–7
- overlap density, 187
- partial molar quantities, 83–84
- particle in a box, 159–62
- Pauli principle, 182
- peptide-induced conformational change of MHC protein, 432–4
- peptides, 450–452
- phase changes, 36–9
- phase diagrams, 37
- phosphorescence, 236–40
- photoelectric effect, 155–6
- plane polarized light, 253
- β -pleated sheets proteins, 58
- polarizability, 279
- polarization, 268, *See also* fluorescence
 - polarization
- polarized light, 253
- potential energy surfaces, 358–64
- preparative centrifugation, 425–7
- principle of detailed balance, 113
- probability density, 158
- processive reaction, 241
- prosthetic groups, 231–3
 - heme, 231–2
- proteasome structure and function, 328–9
- protein–DNA interactions, 318–20
- protein folding, 60–63, 263–6, 452–5
 - dynamics of, 320–322
- protein structure, 52–60, 216–19
 - α -helix, 57
 - β -pleated sheets proteins, 58
 - polypeptide chain, 53
 - primary structure, 216
 - quaternary structure, 218
 - secondary structure, 216
 - tertiary structure, 217
- protein tyrosine phosphatase, 133–7
- proteins, 257–9, 450–452
 - circular dichroism of, 257–9
 - optical rotary dispersion of, 257–9
 - ultraviolet spectra of, 228–30
- P–V isotherm, 8
- quadrupole mass analyzer, 443
- quantum chemistry of biological systems, 194–200
 - global minimum, 196
 - local minima, 196
- quantum dynamics, 358
- quantum mechanical/molecular mechanical (QM/MM) methods, 374
- quantum mechanics, 155–75, *See also* spectroscopy and electronic structure
 - fundamentals of, 155–75
 - particle in a box, 159–62
 - rotational motions, 167–8
 - Schrödinger equation, 158–9
 - tunneling, 165–7

quantum mechanics (*Continued*)

vibrational motions, 162–5

quantum yield, 238

radial distribution function, 180

radioactive decay, 104–5

Raman scattering, 279

Raman spectroscopy, 279–81

advantages, 280

Raoult's law, 82

rate constants, 101

rate constants, temperature dependence of, 108–12

activation energy, 109

standard Gibbs energy of activation, 110

transition state theory, 109

rate laws, determination of, 100–4

first-order kinetics, 102

pseudo-first-order kinetics, 103

second-order kinetics, 103

rates of chemical reactions, *See* reaction rates

reaction enthalpies, 16–18

temperature dependence of, 18–19

reaction field method, 367

reaction mechanisms, 105–8

elementary steps, 105

reaction order, 100

reaction rates, 99–101

near equilibrium, 114–16

RecBCD, helicase activity monitored by

fluorescence, 240–241

relaxation times, 298–300

renaturation, 142–8

residual dipolar coupling (RDC), 303

resonance Raman spectroscopy, 280, 283–6

reversible path, 8

ribozymes, 137–42

ring currents, 296

RNA, 72

RNA folding, 322–5

root-mean-square deviation (RMSD), 370–371

root-mean-square fluctuation (RMSF), 371

rotational motions, 167–8

scattering phenomenon, 279

Schrödinger equation, 158–9

second law of thermodynamics, 24

statement of, 24–6

second-order kinetics, 103

sedimentation coefficient, 421

selection rules, 225

self-consistent-field (SCF) method, 192

simple batch calorimeter, 14

single molecule kinetics, 116–18

singlet, 239

small molecules binding to multiple

identical binding sites, 385–7

sodium dodecyl sulfate–polyacrylamide gel

electrophoresis (SDS-PAGE), 431

spectroscopy, 169–73

basics of, 169–73

constructive interference, 170

destructive interference, 170

spin–lattice relaxation time, 298

spin–spin relaxation time, 298

spin–spin splitting, 296–7

standard electrochemical potentials, 471–2

concentration dependence of, 472–3

standard enthalpy changes, 465, 467

standard Gibbs energy, 465, 467

of activation, 110

standard states, 12–13, 32

state function, 10

statistical effects in ligand binding to

macromolecules, 389–92

statistical entropy, 349–50

statistical mechanics, 333–55

Boltzmann distribution, 338–43

canonical distribution, 348

ensembles, 346–9

fundamentals of, 333–55

helix-coil transition, 350–353

kinetic model of gases, 333–8

molecular partition function, 343–6

statistical entropy, 349–50

steady-state approximation, 123

structure determination with vibrational spectroscopy, 281–3

system, 4

closed system, 4

extensive properties, 4

intensive properties, 4

isolated system, 4

open system, 4

properties, 4

- tandem mass spectrometry (MS/MS), 445
- temperature, 4
 - Celsius temperature scale, 4
 - Kelvin temperature scale, 4
- temperature dependence of Gibbs energy, 35–6
- temperature dependence of the reaction
 - enthalpy, 18–19
- thermal equilibrium, 4
- thermodynamics, 28–9, 77–94
 - applications to biological systems, 43–75
 - biochemical reactions, 43–4
 - chemical equilibria, 87–9
 - chemical potential, 80–83
 - first law of thermodynamics, 9
 - and kinetics, 112–14
 - ionic solutions, 89–93
 - mathematical tools, 77–8
 - Maxwell relations, 78–80
 - osmotic pressure, 85–7
 - partial molar quantities, 83–4
 - second law of thermodynamics, 24–6
 - third law of thermodynamics, 28–9
- third law of thermodynamics, 28–9
- trans conformations, 360
- transient kinetics, 127
- transition state (TS), 196
- transition state theory, 109, 112
- translin, 434
- transverse relaxation optimized spectroscopy (TROSY), 305
- tunneling, 165–7
- ultracentrifuge analysis of protein–DNA
 - interactions, 434–5
- ultraviolet spectra of proteins, 228–30
- umbrella sampling, 375
- van der Waals repulsion energy, 362–3
- velocity sedimentation, 422–4
- velocity Verlet algorithm, 368–9
- vibrational motions, 162–5
 - normal modes, 163
- vibrational spectroscopy, structure
 - determination with, 281–3
- viscosity, 428–9
 - intrinsic viscosity, 428–9
- work, 6–9
- work function, 155
- X-ray absorption spectroscopy, 236
- X-ray crystallography, 205–24, 282
 - anomalous scattering, 210
 - coherent scattering, 210
 - elastic scattering, 207
 - enzyme catalysis, 219–22
 - ‘hammerhead’ ribozyme, 215
 - inelastic scattering, 207
 - neutron diffraction, 212
 - nucleic acid structure, 213–15
 - protein secondary structure determined
 - by, 282
 - protein structure, 216–19
 - scattering of X-rays by a crystal, 206–8
 - structure determination, 208–12
- zinc finger proteins, DNA interaction with, 266–7

Methods of Biochemical Analysis

Biochemical analysis is a rapidly expanding field and is a key component of modern drug discovery and research. *Methods of Biochemical Analysis* provides a periodic and authoritative review of the latest achievements in biochemical analysis. Founded in 1954 by Professor David Glick, *Methods of Biochemical Analysis* provides a timely review of the latest developments in the field.

- Vol. 1. Methods of Biochemical Analysis
- Vol. 2. Methods of Biochemical Analysis
- Vol. 3. Methods of Biochemical Analysis
- Vol. 4. Methods of Biochemical Analysis
- Vol. 5. Methods of Biochemical Analysis
- Vol. 6. Methods of Biochemical Analysis
- Vol. 7. Methods of Biochemical Analysis
- Vol. 8. Methods of Biochemical Analysis
- Vol. 9. Methods of Biochemical Analysis
- Vol. 10. Methods of Biochemical Analysis
- Vol. 11. Methods of Biochemical Analysis
- Vol. 12. Methods of Biochemical Analysis
- Vol. 13. Methods of Biochemical Analysis
- Vol. 14. Methods of Biochemical Analysis
- Vol. 15. Methods of Biochemical Analysis
- Vol. 16. Methods of Biochemical Analysis
- Vol. 17. Methods of Biochemical Analysis
- Vol. 18. Methods of Biochemical Analysis
- Vol. 19. Methods of Biochemical Analysis
- Vol. 20. Methods of Biochemical Analysis
- Vol. 21. Methods of Biochemical Analysis
- Vol. 22. Methods of Biochemical Analysis
- Vol. 23. Methods of Biochemical Analysis
- Vol. 24. Methods of Biochemical Analysis
- Vol. 25. Methods of Biochemical Analysis
- Vol. 26. Methods of Biochemical Analysis
- Vol. 27. Methods of Biochemical Analysis
- Vol. 28. Methods of Biochemical Analysis
- Vol. 29. Methods of Biochemical Analysis
- Vol. 30. Methods of Biochemical Analysis
- Vol. 31. Methods of Biochemical Analysis
- Vol. 32. Methods of Biochemical Analysis
- Vol. 33. Methods of Biochemical Analysis
- Vol. 34. Methods of Biochemical Analysis: Biomedical Applications of Mass Spectrometry
- Vol. 35. Methods of Biochemical Analysis: Protein Structure Determination,
- Vol. 36. Methods of Biochemical Analysis: Bioanalytical Applications of Enzymes

- Vol. 37. Methods of Biochemical Analysis: Bioanalytical Instrumentation
- Vol. 38. HPLC in Enzymatic Analysis, Second Edition,
- Vol. 39. Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins
- Vol. 43. Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins, Second Edition
- Vol. 44. Structural Bioinformatics
- Vol. 45. Industrial Proteomics: Applications for Biotechnology and Pharmaceuticals
- Vol. 47. Green Fluorescent Protein: Properties, Applications, and Protocols, Second Edition
- Vol. 48. Biochemical Thermodynamics: Applications of Mathematica
- Vol. 49. Microbial Proteomics: Functional Biology of Whole Organisms
- Vol. 50. Physical Chemistry for the Biological Sciences
- Vol. 51. Drug-DNA Interactions: Structures and Spectra
- Vol. 52. Introduction to Proteomics: Principles and Applications
- Vol. 53. Enzyme Kinetics: Rapid-Equilibrium Enzyme Kinetics
- Vol. 54. Protein Purification: Principles, High Resolution Methods, and Applications, Third Edition
- Vol. 55. Physical Chemistry for the Biological Sciences, Second Edition

WILEY END USER LICENSE AGREEMENT

Go to www.wiley.com/go/eula to access Wiley's ebook EULA.