PHYSICAL BIOCHEMISTRY: PRINCIPLES AND APPLICATIONS

Second Edition

David Sheehan

Department of Biochemistry
University College Cork
Ireland

WILEY-BLACKWELL
A John Wiley & Sons, Ltd, Publication

PHYSICAL BIOCHEMISTRY

PHYSICAL BIOCHEMISTRY: PRINCIPLES AND APPLICATIONS

Second Edition

David Sheehan

Department of Biochemistry
University College Cork
Ireland

WILEY-BLACKWELL
A John Wiley & Sons, Ltd, Publication

This edition first published 2009, © 2009 John Wiley & Sons Ltd

Wiley-Blackwell is an imprint of John Wiley & Sons, formed by the merger of Wiley's global Scientific, Technical and Medical business with Blackwell Publishing.

Registered office: John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK

Other Editorial Offices:
9600 Garsington Road, Oxford, OX4 2DQ, UK

111 River Street, Hoboken, NJ 07030-5774, USA

For details of our global editorial offices, for customer services and for information about how to apply for permission to reuse the copyright material in this book please see our website at www.wiley.com/wiley-blackwell

The right of the author to be identified as the author of this work has been asserted in accordance with the Copyright, Designs and Patents Act 1988.

All rights reserved. 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 or otherwise, except as permitted by the UK Copyright, Designs and Patents Act 1988, without the prior permission of the publisher.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

Designations used by companies to distinguish their products are often claimed as trademarks. All brand names and product names used in this book are trade names, service marks, trademarks or registered trademarks of their respective owners. The publisher is not associated with any product or vendor mentioned in this book. This publication is designed to provide accurate and authoritative information in regard to the subject matter covered. It is sold on the understanding that the publisher is not engaged in rendering professional services. If professional advice or other expert assistance is required, the services of a competent professional should be sought.

Library of Congress Cataloging-in-Publication Data

Sheehan, David, 1958–
Physical biochemistry : principles and applications / David Sheehan. – 2nd ed.
p. ; cm.
Includes bibliographical references and index.
ISBN 978-0-470-85602-4 (hb) – ISBN 978-0-470-85603-1 (pb) 1. Physical biochemistry. I. Title.
[DNLM: 1. Biophysics. 2. Biochemistry. 3. Chemistry, Phisical. QT 34 S541p 2008]

QD476.2.S42 2009
572′.43–dc22 2008046672

ISBNs: 9780470856024 (HB)
9780470856031 (PB)

A catalogue record for this book is available from the British Library.

Set in 9/11pt Times by Aptara Inc., New Delhi, India
Printed in Singapore by Fabulous Printing Pte Ltd

First Impression 2009

To the memory of my father, Patrick Sheehan (1917–2003)

Contents

Preface **xv**

Chapter 1 Introduction **1**

1.1 Special Chemical Requirements of Biomolecules 1
1.2 Factors Affecting Analyte Structure and Stability 2
  1.2.1 pH Effects 3
  1.2.2 Temperature Effects 3
  1.2.3 Effects of Solvent Polarity 5
1.3 Buffering Systems Used in Biochemistry 6
  1.3.1 How Does a Buffer Work? 6
  1.3.2 Some Common Buffers 7
  1.3.3 Additional Components Often Used in Buffers 7
1.4 Quantitation, Units and Data Handling 7
  1.4.1 Units Used in the Text 7
  1.4.2 Quantification of Protein and Biological Activity 8
1.5 The Worldwide Web as a Resource in Physical Biochemistry 8
  1.5.1 The Worldwide Web 8
  1.5.2 Web-Based Resources for Physical Biochemistry 9
1.6 Objectives of this Volume 9
References 10

Chapter 2 Chromatography **11**

2.1 Principles of Chromatography 11
  2.1.1 The Partition Coefficient 11
  2.1.2 Phase Systems Used in Biochemistry 12
  2.1.3 Liquid Chromatography 12
  2.1.4 Gas Chromatography 13
2.2 Performance Parameters Used in Chromatography 14
  2.2.1 Retention 14
  2.2.2 Resolution 15
  2.2.3 Physical Basis of Peak Broadening 15
  2.2.4 Plate Height Equation 15
  2.2.5 Capacity Factor 19
  2.2.6 Peak Symmetry 19
  2.2.7 Significance of Performance Criteria in Chromatography 20
2.3 Chromatography Equipment 20
  2.3.1 Outline of Standard System Used 20
  2.3.2 Components of Chromatography System 20
  2.3.3 Stationary Phases Used 20
  2.3.4 Elution 21
2.4 Modes of Chromatography 22
  2.4.1 Ion Exchange 22
  2.4.2 Gel Filtration 25
  2.4.3 Reversed Phase 28
  2.4.4 Hydrophobic Interaction 29
  2.4.5 Affinity 31

2.4.6 Immobilized Metal Affinity Chromatography 35
2.4.7 Hydroxyapatite 37
2.5 Open Column Chromatography 37
2.5.1 Equipment Used 37
2.5.2 Industrial Scale Chromatography of Proteins 39
2.6 High Performance Liquid Chromatography (HPLC) 40
2.6.1 Equipment Used 40
2.6.2 Stationary Phases in HPLC 41
2.6.3 Liquid Phases in HPLC 42
2.6.4 Two Dimensional HPLC 42
2.7 Fast Protein Liquid Chromatography 43
2.7.1 Equipment Used 43
2.7.2 Comparison with HPLC 44
2.8 Perfusion Chromatography 44
2.8.1 Theory of Perfusion Chromatography 44
2.8.2 Practice of Perfusion Chromatography 45
2.9 Membrane-Based Chromatography Systems 45
2.9.1 Theoretical Basis 45
2.9.2 Applications of Membrane-Based Separations 46
2.10 Chromatography of a Sample Protein 47
2.10.1 Designing a Purification Protocol 47
2.10.2 Ion Exchange Chromatography of a Sample Protein: Glutathione Transferases 48
2.10.3 HPLC of Peptides From Glutathione Transferases 50
References 50

Chapter 3 Spectroscopic Techniques 53

3.1 The Nature of Light 53
3.1.1 A Brief History of the Theories of Light 53
3.1.2 Wave-Particle Duality Theory of Light 55
3.2 The Electromagnetic Spectrum 55
3.2.1 The Electromagnetic Spectrum 55
3.2.2 Transitions in Spectroscopy 56
3.3 Ultraviolet/Visible Absorption Spectroscopy 58
3.3.1 Physical Basis 58
3.3.2 Equipment Used in Absorption Spectroscopy 61
3.3.3 Applications of Absorption Spectroscopy 62
3.4 Fluorescence Spectroscopy 64
3.4.1 Physical Basis of Fluorescence and Related Phenomena 64
3.4.2 Measurement of Fluorescence and Chemiluminescence 68
3.4.3 External Quenching of Fluorescence 69
3.4.4 Uses of Fluorescence in Binding Studies 72
3.4.5 Protein Folding Studies 73
3.4.6 Resonance Energy Transfer 73
3.4.7 Applications of Fluorescence in Cell Biology 75
3.5 Spectroscopic Techniques Using Plane-Polarized Light 77
3.5.1 Polarized Light 77
3.5.2 Chirality in Biomolecules 78
3.5.3 Circular Dichroism (CD) 79
3.5.4 Equipment Used in CD 80
3.5.5 CD of Biopolymers 81

3.5.6 Linear Dichroism (LD) 83
3.5.7 LD of Biomolecules 83
3.6 Infrared Spectroscopy 84
3.6.1 Physical Basis of Infrared Spectroscopy 84
3.6.2 Equipment Used in Infrared Spectroscopy 86
3.6.3 Uses of Infrared Spectroscopy in Structure Determination 86
3.6.4 Fourier Transform Infrared Spectroscopy 87
3.6.5 Raman Infrared Spectroscopy 90
3.7 Nuclear Magnetic Resonance (NMR) Spectroscopy 91
3.7.1 Physical Basis of NMR Spectroscopy 91
3.7.2 Effect of Atomic Identity on NMR 93
3.7.3 The Chemical Shift 93
3.7.4 Spin Coupling in NMR 94
3.7.5 Measurement of NMR Spectra 95
3.8 Electron Spin Resonance (ESR) Spectroscopy 96
3.8.1 Physical Basis of ESR Spectroscopy 96
3.8.2 Measurement of ESR Spectra 98
3.8.3 Uses of ESR Spectroscopy in Biochemistry 99
3.9 Lasers 99
3.9.1 Origin of Laser Beams 100
3.9.2 Some Uses of Laser Beams 100
3.10 Surface Plasmon Resonance 103
3.10.1 Equipment Used in SPR 105
3.10.2 Use of SPR in Measurement of Adsorption Kinetics 107
References 110

Chapter 4 Mass Spectrometry 113

4.1 Principles of Mass Spectrometry 113
4.1.1 Physical Basis 113
4.1.2 Overview of MS Experiment 115
4.1.3 Ionization Modes 118
4.1.4 Equipment Used in MS Analysis 122
4.2 Mass Spectrometry of Proteins/Peptides 125
4.2.1 Sample Preparation 125
4.2.2 MS Modes Used in the Study of Proteins/Peptides 125
4.2.3 Fragmentation of Proteins/Peptides in MS Systems 125
4.3 Interfacing MS With other Methods 127
4.3.1 MS/MS 127
4.3.2 LC/MS 127
4.3.3 GC/MS 128
4.3.4 Electrophoresis/MS 129
4.4 Uses of Mass Spectrometry in Biochemistry 129
4.4.1 MS and Microheterogeneity in Proteins 130
4.4.2 Confirmation and Analysis of Peptide Synthesis 133
4.4.3 Peptide Mapping 133
4.4.4 Post-Translational Modification Analysis of Proteins 133
4.4.5 Determination of Protein Disulfide Patterns 133
4.4.6 Protein Sequencing by MS 136
4.4.7 Studies on Enzymes 139
4.4.8 Analysis of DNA Components 139
References 143

Chapter 5 Electrophoresis 147

5.1 Principles of Electrophoresis 147
  5.1.1 Physical Basis 147
  5.1.2 Historical Development of Electrophoresis 148
  5.1.3 Gel Electrophoresis 149
5.2 Nondenaturing Electrophoresis 153
  5.2.1 Polyacrylamide Nondenaturing Electrophoresis 153
  5.2.2 Protein Mass Determination by Nondenaturing Electrophoresis 153
  5.2.3 Activity Staining 153
  5.2.4 Zymograms 155
5.3 Denaturing Electrophoresis 155
  5.3.1 SDS Polyacrylamide Gel Electrophoresis 155
  5.3.2 SDS Polyacrylamide Gel Electrophoresis in Reducing Conditions 157
  5.3.3 Chemical Crosslinking of Proteins – Quaternary Structure 159
  5.3.4 Urea Electrophoresis 160
5.4 Electrophoresis in DNA Sequencing 161
  5.4.1 Sanger Dideoxynucleotide Sequencing of DNA 161
  5.4.2 Sequencing of DNA 161
  5.4.3 Footprinting of DNA 165
  5.4.4 Single Strand Conformation Polymorphism Analysis of DNA 165
5.5 Isoelectric Focusing (IEF) 166
  5.5.1 Ampholyte Structure 167
  5.5.2 Isoelectric Focusing 170
  5.5.3 Titration Curve Analysis 170
  5.5.4 Chromatofocusing 170
5.6 Immunoelectrophoresis 172
  5.6.1 Dot Blotting and Immunodiffusion Tests with Antibodies 172
  5.6.2 Zone Electrophoresis/Immunodiffusion Immunoelectrophoresis 174
  5.6.3 Rocket Immunoelectrophoresis 174
  5.6.4 Counter Immunoelectrophoresis 176
  5.6.5 Crossed Immunoelectrophoresis (CIE) 176
5.7 Agarose Gel Electrophoresis of Nucleic Acids 177
  5.7.1 Formation of an Agarose Gel 177
  5.7.2 Equipment for Agarose Gel Electrophoresis 177
  5.7.3 Agarose Gel Electrophoresis of DNA and RNA 177
  5.7.4 Detection of DNA and RNA in Gels 179
5.8 Pulsed Field Gel Electrophoresis 179
  5.8.1 Physical Basis of Pulsed Field Gel Electrophoresis 179
  5.8.2 Equipment Used for Pulsed Field Gel Electrophoresis 181
  5.8.3 Applications of Pulsed Field Gel Electrophoresis 182
5.9 Capillary Electrophoresis 183
  5.9.1 Physical Basis of Capillary Electrophoresis 183
  5.9.2 Equipment Used in Capillary Electrophoresis 188
  5.9.3 Variety of Formats in Capillary Electrophoresis 188
5.10 Electroblotting Procedures 190
  5.10.1 Equipment Used in Electroblotting 190
  5.10.2 Western Blotting 190
  5.10.3 Southern Blotting of DNA 192
  5.10.4 Northern Blotting of RNA 194
  5.10.5 Blotting as a Preparative Procedure for Polypeptides 195

5.11 Electroporation 196
5.11.1 Transformation of Cells 196
5.11.2 Physical Basis of Electroporation 196
References 196

Chapter 6 Three-Dimensional Structure Determination of Macromolecules 199

6.1 The Protein-Folding Problem 199
6.1.1 Proteins are only Marginally Stable 200
6.1.2 Protein Folding as a Two-State Process 203
6.1.3 Protein-Folding Pathways 204
6.1.4 Chaperones 206
6.2 Structure Determination by NMR 212
6.2.1 Relaxation in One-Dimensional NMR 212
6.2.2 The Nuclear Overhauser Effect (NOE) 214
6.2.3 Correlation Spectroscopy (COSY) 215
6.2.4 Nuclear Overhauser Effect Spectroscopy (NOESY) 217
6.2.5 Sequential Assignment and Structure Elucidation 218
6.2.6 Multi-Dimensional NMR 221
6.2.7 Other Applications of Multi-Dimensional NMR 221
6.2.8 Limitations and Advantages of Multi-Dimensional NMR 224
6.3 Crystallization of Biomacromolecules 225
6.3.1 What are Crystals? 226
6.3.2 Symmetry in Crystals 226
6.3.3 Physical Basis of Crystallization 228
6.3.4 Crystallization Methods 231
6.3.5 Mounting Crystals for Diffraction 233
6.4 X-Ray Diffraction by Crystals 235
6.4.1 X-Rays 235
6.4.2 Diffraction of X-Rays by Crystals 235
6.4.3 Bragg's Law 236
6.4.4 Reciprocal Space 238
6.5 Calculation of Electron Density Maps 239
6.5.1 Calculation of Structure Factors 240
6.5.2 Information Available from the Overall Diffraction Pattern 241
6.5.3 The Phase Problem 241
6.5.4 Isomorphous Replacement 242
6.5.5 Molecular Replacement 244
6.5.6 Anomalous Scattering 245
6.5.7 Calculation of Electron Density Map 250
6.5.8 Refinement of Structure 251
6.5.9 Synchrotron Sources 253
6.6 Other Diffraction Methods 254
6.6.1 Neutron Diffraction 254
6.6.2 Electron Diffraction 254
6.7 Comparison of X-Ray Crystallography with Multi-Dimensional NMR 255
6.7.1 Crystallography and NMR are Complementary Techniques 255
6.7.2 Different Attributes of Crystallography- and NMR-derived Structures 256
6.8 Structural Databases 257
6.8.1 The Protein Database 257
6.8.2 Finding a Protein Structure in the Database 257
References 259

Chapter 7 Hydrodynamic Methods **263**

7.1 Viscosity 263
7.1.1 Definition of Viscosity 263
7.1.2 Measurement of Viscosity 264
7.1.3 Specific and Intrinsic Viscosity 265
7.1.4 Dependence of Viscosity on Characteristics of Solute 266
7.2 Sedimentation 266
7.2.1 Physical Basis of Centrifugation 266
7.2.2 The Svedberg Equation 268
7.2.3 Equipment Used in Centrifugation 269
7.2.4 Subcellular Fractionation 272
7.2.5 Density Gradient Centrifugation 273
7.2.6 Analytical Ultracentrifugation 274
7.2.7 Sedimentation Velocity Analysis 274
7.2.8 Sedimentation Equilibrium Analysis 276
7.3 Methods for Varying Buffer Conditions 279
7.3.1 Ultrafiltration 281
7.3.2 Dialysis 282
7.3.3 Precipitation 284
7.4 Flow Cytometry 286
7.4.1 Flow Cytometer Design 286
7.4.2 Cell Sorting 287
7.4.3 Detection Strategies in Flow Cytometry 288
7.4.4 Parameters Measurable by Flow Cytometry 288
References and Further Reading 290

Chapter 8 Biocalorimetry **293**

8.1 The Main Thermodynamic Parameters 293
8.1.1 Activation Energy of Reactions 293
8.1.2 Enthalpy 295
8.1.3 Entropy 295
8.1.4 Free Energy 296
8.2 Isothermal Titration Calorimetry 296
8.2.1 Design of an Isothermal Titration Calorimetry Experiment 296
8.2.2 ITC in Binding Experiments 297
8.2.3 Changes in Heat Capacity Determined by Isothermal Titration Calorimetry 297
8.3 Differential Scanning Calorimetry 300
8.3.1 Outline Design of a Differential Scanning Calorimetry Experiment 300
8.3.2 Applications of Differential Scanning Calorimetry 301
8.4 Determination of Thermodynamic Parameters by Non-Calorimetric Means 301
8.4.1 Equilibrium Constants 301
References 302

Chapter 9 Bioinformatics **305**

9.1 Overview of Bioinformatics 305
9.2 Sequence Databases 309
9.2.1 Nucleotide Sequence Databases 309

9.2.2 Protein Sequence Databases 309
9.2.3 Genome Databases 309
9.2.4 Expressed Sequence Tag Databases 315
9.2.5 Single Nucleotide Polymorphism (SNP) Database 315
9.3 Tools for Analysis of Primary Structures 315
9.3.1 BLAST Programs 317
9.3.2 FastA 318
9.3.3 Clustal W 318
9.3.4 Hydropathy Plots 319
9.3.5 Predicting Secondary Structure 323
9.3.6 Identifying Protein Families 325
9.4 Tertiary Structure Databases 327
9.4.1 Cambridge Database 329
9.4.2 Protein Databank (PDB) 329
9.4.3 Specialist Structural Databases 331
9.5 Programs for Analysis and Visualization of Tertiary Structure Databases 334
9.5.1 Ras Mol/Ras Top 334
9.5.2 Protein Explorer 334
9.5.3 Py MOL 334
9.5.4 Web Mol 336
9.5.5 Swiss-Pdb Viewer 336
9.6 Homology Modelling 336
9.6.1 Modelling Proteins from Known Homologous Structures 340
9.6.2 Automated Modelling 342
9.6.3 Applications of Homology Modelling to Drug Discovery 346
References 346

Chapter 10 Proteomics **349**

10.1 Electrophoresis in Proteomics 349
10.1.1 Two-Dimensional SDS PAGE 350
10.1.2 Basis of 2-D SDS PAGE 350
10.1.3 Equipment Used in 2-D SDS PAGE 350
10.1.4 Analysis of Cell Proteins 351
10.1.5 Free Flow Electrophoresis 353
10.1.6 Blue Native Gel Electrophoresis 354
10.1.7 Other Electrophoresis Methods Used in Proteomics 355
10.2 Mass Spectrometry in Proteomics 355
10.2.1 Tagging Methodologies Used in MS Proteomics 355
10.2.2 Isotope-Coded Affinity Tagging (ICAT) for Cysteine-Containing Proteins 357
10.2.3 Tagging of N- and C-Termini 358
10.2.4 Tagging for Tandem MS 359
10.3 Chip Technologies in Proteomics 359
10.3.1 Microarrays 359
10.3.2 Protein Biochips 362
10.3.3 SELDI-TOF MS on Protein Chips 362
10.4 Post-Translational Modification Proteomics 366
10.4.1 Proteolysis 366
10.4.2 Glycosylation 367
10.4.3 Oxidation 372
10.4.4 Protein Disulfides 374

10.4.5 The Phosphoproteome 376
Further Reading 379
References 380

Appendix 1 SI Units 381

Appendix 2 The Fourier Transform 383

Index 387

Preface

The first edition of *Physical Biochemistry: Principles and Applications* set out to describe the physical basis and some examples of applications of key physically-based techniques used in Biochemistry and other areas of molecular life science research. In the last decade there has been a noticeable renaissance in some traditional techniques such as X-ray diffraction, ultracentrifugation and electrophoresis in a variety of formats. In the same time-frame 'hyphenated' techniques (e.g. LC-tandem MS) have become much more mainstream and some instrumentation has become available in desktop formats making quite sophisticated analysis possible even in the nonspecialist lab. The emphasis was on a largely nonmathematical treatment at a level appropriate to students in the penultimate year of a Biochemistry course with a view to making these techniques comprehensible and accessible to a level intermediate between a general Biochemistry textbook and a specialist text. The feedback I have had from many readers is that this goal was largely achieved.

My task with this second edition was to retain as much as possible of the description of physical principles whilst updating and integrating new material into a reasonably compact volume. The first edition was strongly influenced by the effects of the then-recent completion of the Human and other large-scale genome sequencing projects. At the time Proteomics and other '-omics' technologies were still relatively new paradigms and bioinformatics approaches largely the province of the specialist. It is remarkable how quickly these technologies have now become embedded in many areas of biochemical research so that they are now perceived as part of the mainstream. In the second edition I have dedicated new chapters to proteomics and bioinformatics, respectively, to reflect this changed situation and to emphasize how interconnected physical and computational techniques have now become.

The second edition required choices to be made and these are inevitably influenced by one's perception of what is needed and useful for students to know at the start of their scientific journey into molecular life science. In making these choices I have continued to be guided by what I perceive to be the most generally-used and helpful techniques but I accept that specialists in one or more technique may disagree.

In preparing this second edition I had help from many of the colleagues listed in the preface to the First Edition. In addition, I must thank Dr Rebecca Green, School of Pharmaceutical Sciences, University of Nottingham, UK, for her invaluable comments on surface plasmon resonance (Chapter 3). I am also indebted to the excellent staff at Wiley's especially my editors, Celia Carden and Fiona Woods for their unfailing encouragement and understanding. However, any errors in the text are my own. I earnestly hope that the reader will find something interesting and thought-provoking in this volume and be encouraged to explore these very powerful approaches in their work.

Prof David Sheehan
University College Cork
June 2008

Chapter 1

Introduction

Objectives

After completing this chapter you should be familiar with:

- *The* **special chemical conditions** *often required by biomolecules.*
- *Importance of the use of* **buffers** *in the study of physical phenomena in biochemistry.*
- **Quantification** *of physical phenomena.*
- **Objectives** *of this volume.*

This volume describes a range of physical techniques which are now widely used in the study both of biomolecules and of processes in which they are involved. There will be a strong emphasis throughout on *biomacromolecules* such as proteins and nucleic acids as well as on *macromolecular complexes* of which they are components (e.g. biological membranes, ribosomes, chromosomes). This is because such chemical entities are particularly crucial to the correct functioning of living cells and present specific analytical problems compared to simpler biomolecules such as monosaccharides or dipeptides. Biophysical techniques, give detailed information offering insights into the structure, dynamics and interactions of biomacromolecules.

Life scientists in general and biochemists in particular have devoted much effort during the last century to elucidation of the relationship between structure and function and to understanding how biological processes happen and are controlled. Major progress has been made using *chemical* and *biological* techniques which, for example, have contributed to the development of the science of *molecular biology*. However, in the last decade *physical* techniques which complement these other approaches have seen major development and these now promise even greater insight into the molecules and processes which allow the living cell to survive. For example, a major focus of life science research currently is the *proteome* as distinct from the *genome*. This has emphasized the need to be able to study the highly-individual structures of biomacromolecules such as proteins to understand more fully their particular contribution to the biology of the cell. For the foreseeable future, these techniques are likely to impact to a greater or lesser extent on the activities of most life scientists. This text attempts to survey the main physical techniques and to describe how they can contribute to our knowledge of biological systems and processes. We will set the scene for this by first looking at the particular analytical problems posed by biomolecules.

1.1 SPECIAL CHEMICAL REQUIREMENTS OF BIOMOLECULES

The tens of thousands of biomolecules encountered in living cells may be classified into two general groups. Biomacromolecules (e.g. proteins; nucleic acids) are characterized by high molecular mass (denoted throughout this text as *relative molecular mass*, M_r) and are generally unstable under extreme chemical conditions where they may lose structure or break down into their chemical building blocks. Low molecular weight molecules are smaller and more chemically robust (e.g. amino acids; nucleotides; fatty acids). Within each group there is displayed a wide range of water-solubility, chemical composition and reactivity which is determined by complex interactions between physicochemical attributes of the biomolecule and solvent. These attributes are the main focus for the techniques described in this volume and reflect the highly individual function which each molecule performs in the cell (Tables 1.1 and 1.2).

Notwithstanding the great range of form and structure, we can nonetheless recognize certain attributes as common to all biomolecules. The first and most obvious is that all of these molecules are produced in living cells *under mild chemical conditions* of temperature, pressure and pH. Biomacromolecules are built up from simpler building block molecules by covalent bonds formed usually with the elimination of water. Moreover, *biomolecules are continuously synthesized and degraded* in cells in a highly regulated manner. It follows from this that many biomolecules are especially sensitive to extremes of temperature and pH which may present a problem in their handling prior to and during

---

Physical Biochemistry: Principles and Applications, Second Edition David Sheehan
© 2009 John Wiley & Sons, Ltd

Table 1.1. Some important physical attributes of biomolecules amenable to study by biophysical techniques

Physical attribute	Technique
Mass	MS (Chapters 4 and 10) Electrophoresis (Chapters 5 and 10) Gel filtration (Chapter 2)
Volume/density	Gel filtration (Chapter 2) Centrifugation (Chapter 7) Pulsed field gel electrophoresis (Chapter 5)
Charge	Ion exchange chromatography (Chapter 2) Electrophoresis/chromatofocusing (Chapters 5 and 10) MS (Chapters 4 and 10)
Shape	Chromatography (Chapter 2) Electrophoresis (Chapters 5 and 10) Crystallization (Chapter 6) Centrifugation (Chapter 7)
Energy	Spectroscopy (Chapter 3) Biocalorimetry (Chapter 8)

Table 1.2. Some important chemical attributes of biomolecules which may be used in study by biophysical techniques

Chemical attribute	Technique
Composition	MS (Chapters 4 and 10) Spectroscopy (Chapter 3)
Molecular structure	Spectroscopy (Chapter 3) Crystallization (Chapter 6)
Covalent bonds	MS (Chapters 4 and 10) Bioinformatics (Chapter 9) Electrophoresis (Chapters 5 and 10)
Noncovalent bonds	Chromatography (Chapter 2) Electrophoresis (Chapters 5 and 10) Spectroscopy (Chapter 3)
Native/denatured structure	Chromatography (Chapter 2) Electrophoresis (Chapter 5) Proteomics (Chapter 10)
Solubility	Chromatography (Chapter 2) Crystallography (Chapter 6) Precipitation (Chapter 7)
Complex formation	Electrophoresis (Chapter 5) MS (Chapters 4 and 10) Spectroscopy (Chapter 3) Crystallography (Chapter 6)

any biophysical analysis. Since biomolecules result from a long process of biological evolution during which they have been selected to perform highly specific functions, a very close relationship has arisen between chemical structure and function. This means that, even at pH and temperature values under which the molecule may not be destroyed, it may function suboptimally or not at all.

These facts impose limitations on the chemical conditions to which biomolecules may be exposed during extraction, purification or analysis. In most of the techniques described in this volume, sample analytes are exposed to a specific set of chemical conditions by being dissolved in a solution of defined composition. Whilst other components in the solution may also be important in individual cases as will be discussed below (Section 1.3.2), three main variables govern the makeup of this solution which are discussed in more detail in the following section.

1.2 FACTORS AFFECTING ANALYTE STRUCTURE AND STABILITY

In practice, most of the biophysical procedures described in this volume use conditions which have been optimized over many years for thousands of different samples. These robust conditions will normally maintain the sample in a defined structural form facilitating its separation and/or analysis. However, some procedures (e.g. chromatography, capillary electrophoresis, crystallization) may require case-by-case optimization of conditions. Before embarking on a detailed analysis of a biomolecule using biophysical techniques it is often useful to know something about the stability of the sample to chemical variables, especially pH, temperature and solvent polarity. This knowledge can help us to design a suitable solvent or set of chemical conditions which will maximize the stability of the analyte for the duration of the experiment and may also help us to explain unexpected results. For example, we sometimes find loss of enzyme activity during column chromatography which may be partly explained by the chemical conditions experienced by the protein during the experiment. Moreover, many of the techniques described in this volume are actually designed to be *suboptimal* and to take advantage of *disruption* of the normal functional structure of the biomolecule to facilitate separation or analysis (e.g. electrophoresis, HPLC, MS).

A good indication of the most stabilizing conditions may often be obtained from knowledge of the biological origin of the biomolecule. It is also wise to assess the structural and functional stability of the analyte over the range of experimental conditions encountered in the experiment during its likely time-span.

We can distinguish two main types of effects as a result of variation in the chemical conditions to which biomolecules are exposed. *Structural effects* reflect often irreversible structural change in the molecule (e.g. protein/nucleic acid denaturation; hydrolysis of covalent bonds between building blocks of which biopolymers are composed). *Functional effects* are frequently more subtle and may be reversible (e.g. deprotonation of chemical groups in the biomolecule resulting in ionization; partial unfolding of proteins). A detailed treatment of these effects on the main classes of biomolecules is outside the scope of the present volume but a working knowledge of the likely effects of these conditions can be very useful in deciding conditions for separation or analytical manipulation.

1.2.1 pH Effects

pH is defined as the negative log of the proton concentration:

$$\mathrm{pH} = -\log[\mathrm{H}^+] \tag{1.1}$$

Because both the H^+ and OH^- concentrations of pure water are 10^{-7} M, this scale runs from a *maximum* of 14 (strongly alkaline) to a minimum of 0 (strongly acidic). As it is a log scale, one unit reflects a *10-fold* change in proton concentration. Most biomacromolecules are *labile* to alkaline or acid-catalyzed hydrolysis at extremes of the pH scale but are generally *stable* in the range 3–10. It is usual to analyse such biopolymers at pH values where they are structurally stable and this may differ slightly for individual biopolymers. For example, proteins normally expressed in lysosomes (pH 4) are quite acid-stable while those from cytosol (pH 7) may be unstable near pH 3. Aqueous solutions in which sample molecules are dissolved usually comprise a *buffer* to prevent changes in pH during the experiment. These are described in more detail in Section 1.3 below.

Many biomolecules are *amphoteric* in aqueous solution that is they can accept or donate protons. Some chemical groups such as inorganic phosphate or acidic amino acid side-chains (e.g. aspartate) can act as *Brønsted acids* and donate protons:

$$\mathrm{AH} \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} \mathrm{A}^- + \mathrm{H}^+ \tag{1.2}$$

Other groups such as the imidazole ring of histidine or amino groups can act as *Brønsted bases* and accept protons:

$$\mathrm{B}^- + \mathrm{H}^+ \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} \mathrm{BH} \tag{1.3}$$

The position of equilibrium in these protonation/deprotonation events may be described by an *equilibrium constant*, K_a:

$$K_a = \frac{k_1}{k_{-1}} = \frac{[\mathrm{A}^-]\cdot[\mathrm{H}^+]}{[\mathrm{AH}]} \tag{1.4}$$

pK_a ($-\log K_a$) is the pH value at which 50% of the acid is protonated and 50% is deprotonated. The *Henderson–Hasselbach equation* describes variation of concentrations of A^- and AH as a function of pH:

$$\mathrm{pH} = \mathrm{p}K_a + \frac{\log[\mathrm{A}^-]}{[\mathrm{AH}]} \tag{1.5}$$

Functional groups present as structural components of biomolecules (e.g. amino acid side-chains of proteins; phosphate groups of nucleotides) will have distinct K_a values which may differ slightly from the value found in other chemical circumstances (e.g. the K_a values of amino acid side-chains in polypeptides differ from those in the free amino acid). Some biomolecules can contain *both* acidic and basic groups within their structure (e.g. proteins) while particular chemical structures found in biomolecules may be *polyprotic*, that is capable of *multiple* ionizations (e.g. phosphate). Such biomolecules may undergo a complex pattern of ionization resulting in varying net charge on the molecule. *pH titration curves* for biomolecules allow us to identify pK_a values (Figure 1.1).

Since protonation-deprotonation effects are responsible for the charges on biomacromolecules which maintain their solubility in water, their solubility is often *lowest* at their *isoelectric point*, pI, the pH value at which the molecule has no net charge. These can also be determined by titration using methods described in Chapter 5 (Section 5.5.3; Figure 5.24).

While the pH scale reflects the situation in aqueous solution, many microenvironments encountered in living cells are quite nonpolar (*see* below). Good examples include biological membranes and water-excluding regions of proteins (e.g. some enzyme active sites). In these environments, protonation/deprotonation properties of chemical groups may deviate widely from those observed in aqueous solution. For example, catalytic residues of many enzymes frequently display pK_a values which are perturbed far from those normal for that residue in water-exposed regions of proteins.

1.2.2 Temperature Effects

Three main effects of temperature on biomolecules are important for the biophysical techniques described in this volume. These are effects on structure, chemical reactivity and solubility. Heat can disrupt noncovalent bonds such as

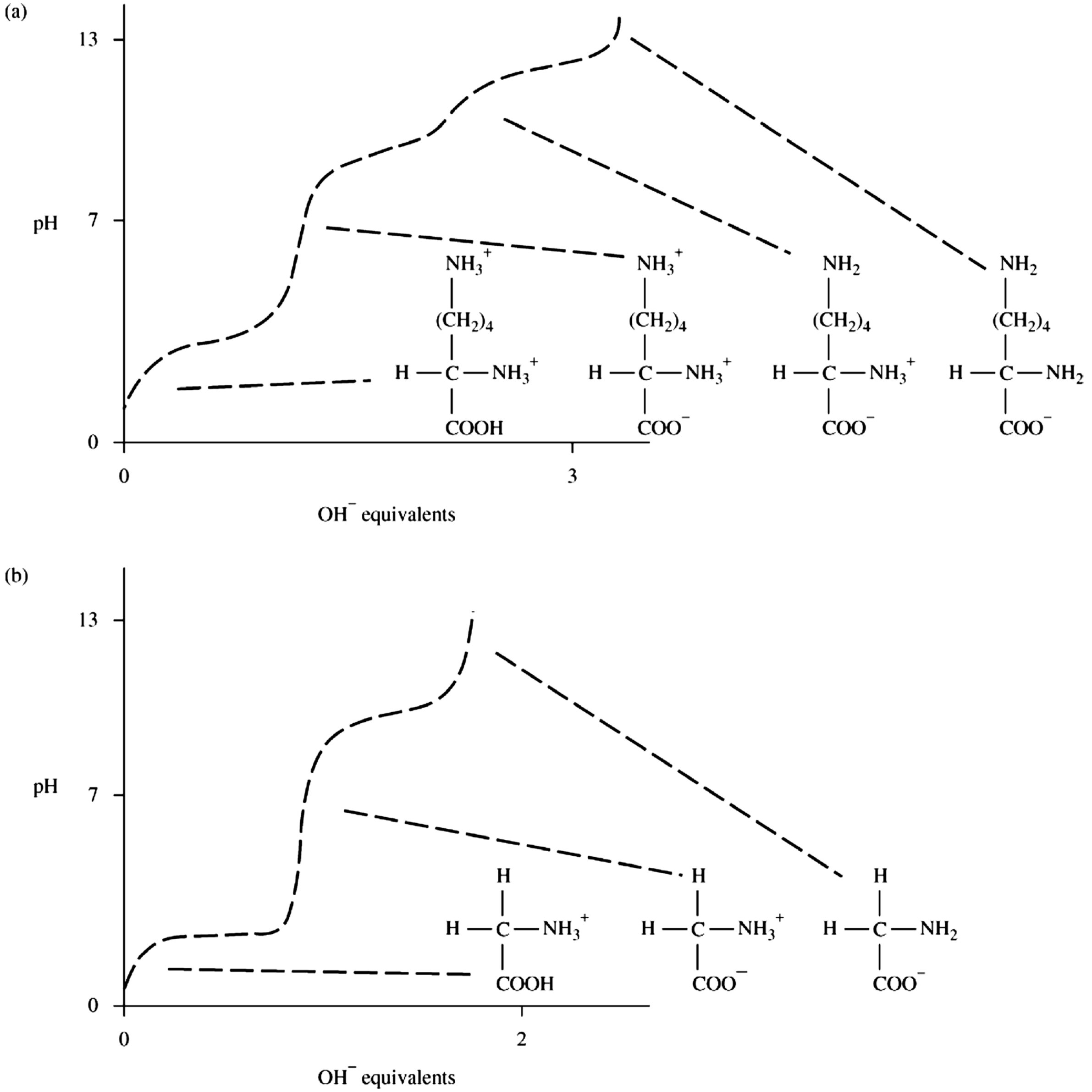

Figure 1.1. pH titration curves, (a) Lysine. Four protonation states are possible for lysine as shown. Three pK_a values are evident from this at pH values of 2.18, 8.95 and 10.53. The pI of lysine is 9.74. (b) Glycine. Note that only three protonation states exist for glycine compared to four for lysine. pK_a values are at pH values of 2.34 and 9.6 which differ slightly from the corresponding ionisations in (a). Glycine has a pI of 5.97

hydrogen bonds which are especially important in the structure of biomacromolecules. This can lead to denaturation of proteins and DNA or to disruption of multimolecular complexes in which they may be involved. Moreover, since covalent bonds linking building block molecules (e.g. peptide bonds; glycosidic bonds; $3', 5'$-phosphodiester bonds) have generally lower bond energies than bonds within such building blocks, extensive heating can result in disintegration of the covalent structure of biomacromolecules. Thus proteins can break down into component peptides or nucleic acids into smaller polynucleotide fragments as a result of exposure to heat.

Secondly, all chemical reaction processes obey the Arrhenius relationship:

$$k = A\,e - E_a/RT \qquad (1.6)$$

where k is the rate constant for the process, A is a constant, E_a is the activation energy of the reaction, R is the Universal gas constant and T is absolute temperature. This relationship arises from large changes in the number of activated molecules available for reaction as a result of change in temperature. The exponential dependence on temperature means that small changes in T can result in large effects on the rate constant, k.

Thirdly, temperature usually increases the solubility of molecules in a solvent as well as the rate of diffusion through the solvent. This is because heat increases the average kinetic energy of solvent molecules. In the case of water, this is accompanied by extensive breakdown of water-water hydrogen bonds which increases the solute capacity of a given volume of water. Thus, for example 8 M urea is soluble at 30 °C while the limit of solubility is closer to 5 M at 4 °C. Kinetic energy effects are also important in situations involving biological membranes because the phospholipid bilayer of which they are composed becomes increasingly fluid at higher temperature.

Temperature is therefore usually tightly controlled during biophysical experiments. In dealing with biomacromolecules in particular, it is generally not possible to use temperatures higher than 80 °C and, in most cases, much lower temperatures are used. Moreover, samples such as proteins or nucleic acids are normally stored under refrigerated conditions to maximize their stability. This is achieved with the aid of liquid nitrogen (−196 °C) or with refrigerators set at −80 or −20 °C. Particular care must be taken in handling crude biological extracts since hydrolases such as proteases and nucleases present in these will be active in the range 18–37 °C and result in extensive degradation of proteins and nucleic acids. This can be avoided by maintaining low temperatures near 4 °C during manipulation of sample and by cooling buffer solutions before dissolving biological samples.

Most biomolecules are optimally active at temperatures similar to those experienced in the biological source from which they were obtained. For example, proteins from thermophilic bacteria are especially heat stable compared to corresponding proteins from mesophilic bacteria while mammalian proteins are optimally active around 37 °C.

1.2.3 Effects of Solvent Polarity

Polarity arises from unequal affinity of atoms bonded together for shared electrons called electronegativity. Apart from fluorine (with an electronegativity value of 4), oxygen is the most electronegative element in the periodic table (electronegativity value of 3.5) leading, for example, to oxygens of –OH groups tending to be partially negatively charged (δ^-) while hydrogens tend to be partially positively charged (δ^+). Water (which is itself polar) interacts *ionically* with polar functional groups such as –OH by hydrogen bonding (Figure 1.2). Organic molecules composed mainly of carbon and hydrogen, however, tend to be nonpolar as these atoms have similar electronegativities (2.5 and 2.1, respectively). In general, polar biomolecules dissolve readily in polar solvents such as water while those which are nonpolar dissolve in nonpolar solvents (e.g. trichloromethane).

Biomolecules lacking strongly electronegative elements such as oxygen and nitrogen and consisting mainly of carbon and hydrogen tend to be principally nonpolar (e.g. fatty acids; sterols; integral membrane proteins). Conversely, those containing oxygen, sulfur and nitrogen tend to be mainly polar (e.g. monosaccharides; nucleotides). Biomacromolecules often contain distinct structural regions some of which may be polar while others may be nonpolar.

Since water is the main biological solvent, most biomolecules (or parts of biomolecules) have been selected by evolution to interact with it in particular ways either by attraction or repulsion. Polar regions strongly *attracted* to water are called *hydrophilic* while nonpolar regions which are *repulsed* by water are called *hydrophobic*. In the living cell, biomolecules adopt a structure determined to a large degree by the extent to which they are hydrophobic/hydrophilic. For example, biological membranes are made up of phospholipid bilayers which spontaneously form when phospholipid molecules are dissolved in water. The polar heads of phospholipids are on the exterior in contact with water while the nonpolar fatty acid components are on the interior of the bilayer protected from water. Cytosolic proteins express hydrophilic groups on their surface whilst folding in such a manner that hydrophobic groups are protected from exposure to water in the interior of the protein (Chapter 6). Membrane-bound proteins such as hormone receptors expose hydrophobic groups to the interior of biological membranes and hydrophilic groups to the exterior.

In extracting, analyzing and purifying biomolecules these intricate structural interactions are often lost which can result in aggregation, precipitation or loss of structure and, hence, of biological activity. If it is desired to retain biological activity we use aqueous solutions to handle largely hydrophilic biomolecules, nonpolar solvents to dissolve mainly hydrophobic samples and detergent solutions for molecules which possess both classes of groups. Many of the individual techniques described in this volume use specific solvent systems of distinct polarity/nonpolarity but it may occasionally be necessary to design individual solvent systems to take account of the requirements of particular biomolecules. Examples include column chromatography

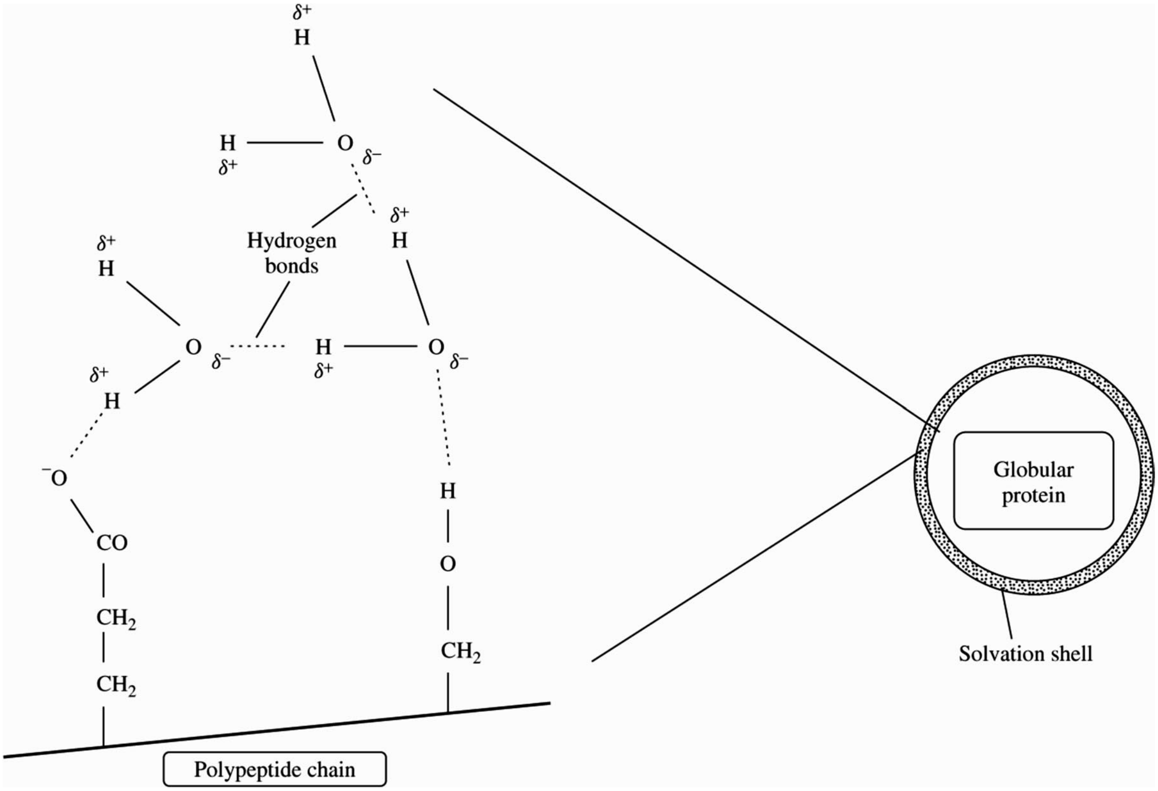

Figure 1.2. Hydrogen bonding. Water molecules hydrogen bond because of partial charge-differences arising from the different electronegativities of oxygen and hydrogen. Hydrogen bonds (dashed lines) are shown both in bulk water and between water and amino acid side-chains of glutamic acid and serine residues of a polypcptidc. Such ionic interactions maintain a solvation shell of water around the surface of globular proteins and other hydrated biomacromolecules

(Chapter 2), spectroscopy (Chapter 3) and capillary electrophoresis (Chapter 5).

1.3 BUFFERING SYSTEMS USED IN BIOCHEMISTRY

A buffer is an aqueous solvent system designed to maintain a given pH. In the context of biochemical work, the main function of buffers is to resist any tendency for pH to rise or drop during the experiment. This can happen during any process which might release or absorb protons from solution such as, for example, during an enzyme-catalyzed reaction or as a result of electrochemical processes such as electrophoresis. A secondary but often crucial role for a buffer is to maximize the stability of biomolecules in solution. Frequently, additional molecules are dissolved in the buffer to help it to do this and these are discussed in more detail below.

1.3.1 How Does a Buffer Work?

Any aqueous solution containing both A^- and AH (Section 1.2 above) is, in principle, capable of resisting change in pH. This is because, if protons are generated in the solution, they can be neutralized by A^-:

$$A^- + H^+ \rightarrow AH \qquad (1.7)$$

Conversely, if alkali is generated in the solution (which would tend to *remove* protons), it can be neutralized by AH:

$$AH + OH^- \rightarrow A^- + H_2O \qquad (1.8)$$

In practice, most buffers consist of mixtures either of a weak acid and its salt or of a weak base and its salt.

Of course, the ability of a buffer to resist change in pH is finite, especially if the number of protons involved is especially large. This limit is represented by the *buffering capacity* of the buffer, β. This is defined as the number of moles of $[H^+]$ which must be added to a liter of the buffer to decrease the pH by one unit. It can be mathematically calculated from the following equation:

$$\beta = \frac{2.3 \cdot K_a \cdot [H^+] \cdot [C]}{(K_a + [H^+])^2} \qquad (1.9)$$

where [C] is the *sum* of the concentrations of A^- and AH. This relationship means that *buffering capacity increases with buffer concentration* so that, for example 100 mM acetate buffer has 50-fold greater buffering capacity than 2 mM. It can be demonstrated experimentally that β reaches a maximum at pH values equal to pK_a which is when the concentrations of A^- and AH are approximately equal. This means that buffers work best at pH values around their pK_a. In practice, most buffers are effective one pH unit above and one below their pK_a so, for example acetate buffers ($pK_a = 4.8$) are useful in the pH range 3.8 to 5.8, although most effective around pH 4.8.

1.3.2 Some Common Buffers

A selection of buffers commonly used in biochemistry is given in Table 1.3. Some of these buffer components are of biological origin (e.g. glycine; histidine; acetate). *Good's buffers* were developed by N.E. Good to facilitate buffering in the pH range 6–10.5. Because of their complicated chemical names, these buffers are more usually known by abbreviations (e.g. Pipes; Hepes; Mops). Inspection of Table 1.3 shows that buffers are available which span the range of interest for physical studies of biomolecules (i.e. pH 3–10). When preparing buffers, it is essential that *both* the concentration and pH are correct since these are the two variables critical to buffering capacity (Equation (1.7)).

A number of problems can arise with particular buffers which can limit their use in specific cases. For example, several buffers interact with divalent metals (e.g. phosphate binds Ca^{2+}; Tris reacts with Cu^{2+} and Ca^{2+}) and should be avoided in cases where this is important to interpretation of the experiment. Tris buffers are especially sensitive to temperature which can result in the same buffer giving a slightly different pH at different temperatures. Phosphate buffers are particularly susceptible to bacterial contamination if stored for long periods of time (although this can be avoided by including a low concentration of sodium azide as a preservative). Some buffer components (e.g. EDTA) may give high absorbance readings which can affect detection during processes such as chromatography. Volatile buffer components (e.g. formic acid; bicarbonate; triethanolamine) can be lost from the buffer over time leading to a gradual change in pH and buffering capacity.

1.3.3 Additional Components Often Used in Buffers

In addition to buffer components such as weak acids/bases and their salts, buffers frequently contain a range of other components of which a selection is shown in Table 1.4. These may be necessary to maintain stability of the biomolecule, to control levels of metal ions, to ensure reducing/oxidizing conditions or to keep the biomolecule dissolved and/or denatured. We will observe that some of the agents tabulated in Table 1.4 are used in more than one of the techniques described in this book and therefore represent generally useful tools for the manipulation of chemical conditions to which biomolecules are exposed.

Table 1.3. Some common buffers used in biochemistry

Buffer	pK_a
Phosphate[a]	0.85, 1.82, 6.68
Histidine[a]	1.82, 5.98, 9.17
Phosphoric acid[a]	1.96, 6.7, 12.3
Formic acid	3.75
Barbituric acid	3.98
Acetic acid	4.8
Pyridine	5.23
Bis TRIS [Bis-(2-hydroxy-ethyl)imino-tris-(hydroxy-methyl)-methane]	6.46
PIPES [1,4-piperazinebis-(ethanesulphonic acid)]	6.8
Imidazole	7.0
BES [*N*, *N*-*B*is(2-hydroxy-ethyl)-2-amino-ethane-sulphonic acid]	7.15
MOPS [2-(*N*-morpholino)propane-sulphonic acid]	7.2
HEPES [*N*-2-hydroxyethyl-piperazine-*N*′-2-ethane-sulphonic acid]	7.55
TRIS [hydroxymethyl)amino-methane]	8.1
TAPS [W-tris(hydroxymethyl)methyl-2-aminopropane sulphonic acid]	8.4
Boric acid	9.39
Ethanolamine	9.44
CAPS [3-(cyclohexylamino)-l-propane-sulphonic acid]	10.4
Methylamine	10.64
Dimelhylamine	10.75
Diethylamine	10.98

[a] There are polyprotic with several pK_a values.

1.4 QUANTITATION, UNITS AND DATA HANDLING

1.4.1 Units Used in the Text

Physical measurements usually result in *quantification* of some property of a molecule or system such as those tabulated in Table 1.1. Various systems of internationally-agreed units have been used historically to record these measurements but, throughout this book, the *Systeme Internationale* (SI) system, the most currently-agreed scientific

Table 1.4. Additional reagents sometimes added to buffers

Chemical	Function
2-Mercaptoethanol	Reducing agent
Dithiothreitol	Reducing agent
Sodium borohydride	Reducing agent
Divalent metals plus O_2	Oxidising agents
Performic acid	Oxidising agent
Leupeptin	Protease inhibitor
Phenylmethyl sulphonyl fluoride (PMSF)	Serine protease inhibitor
Ethylene diamine *NNN′N′* tetra-acetic acid (EDTA)	Metal chelator/ metalloprotease inhibitor
Ethylene glycol-bis(β-aminoethyl ether) *NNN′N′* tetraacetic acid (EGTA)	Calcium chelator
Urea	Denaturing agent
Guanidinium hydrochloride	Denaturing agent
Sodium dodecyl sulphate (SDS)	Anionic detergent
Cetyltrimethyl ammonium chloride	Cationic detergent
3-(3- cholamidopropyl)dimethy lammonio)-1-propane sulphonate (CHAPS)	Zwitterionic detergent
Triton X-100	Nonionic detergent
Digitonin	Nonionic detergent
Protamine K	Binds DNA

quantification system will be used. The main units are tabulated in Appendix 1. This system has the advantage of great *internal consistency* which removes any need for conversion factors. For example, units of distance (meters) readily relate to units of velocity (meters/second).

In addition to SI units, some measurements used are *operational measurements* commonly employed in the life science literature. These are units which have gained wide acceptance in the international science community but which are not strictly part of the SI system. Relative molecular mass (M_r) is expressed as Daltons (Da) or multiples thereof (e.g. kDa) with 12 Da being equivalent to twelve atomic mass units (i.e. the mass of ^{12}C). In the case of nucleic acids, base pairs (bp) or multiples thereof (e.g. kbp, Mbp) are used as units of mass. Interatomic distances such as bond-lengths are generally given as angstroms (Å) with 1 Å corresponding to 10^{-10} m (i.e. 0.1 nm).

Concentrations are given mainly in molarity (1 M solution being Avogadro's Number of molecules dissolved in 1 l of solvent) although occasionally they are expressed as percentages of weight/weight (% w/w) or weight/volume (% w/v). Thus, 10% (w/v) would represent a solution of 10 g per 100 ml while 10% (v/v) would represent a solution of 10 ml per 100 ml.

The most commonly used temperature scale in the text is the Celsius scale although absolute temperatures (in units of Kelvin, K) are specifically referred to by T (e.g. Equation (1.6) above).

1.4.2 Quantification of Protein and Biological Activity

Most of the techniques described in this text are used to separate or analyse biomolecules or mixtures containing them. In carrying out this kind of experimentation it is crucial to know exactly *how much* sample is being applied since most of the systems described are highly loading-sensitive. In the case of pure samples this may not be a problem but many samples encountered in biochemistry may be quite crude and heterogeneous. A common strategy for quantifying such samples is to estimate their *protein content* (e.g. by the *Bradford* method, Figure 3.18, or by one of the other methods mentioned in the bibliography at the end of this chapter) and to load a standard amount of protein. Since the ratio of protein to the other components of the mixture is fixed, this normally ensures uniform loading. Similarly, when quantifying the biological activity of a sample (e.g. enzyme activity, antibody content, antiviral activity) it is often useful to express this as *specific activity that is* units/mg protein. This is a measure which is *independent* both of sample volume and sample concentration.

The majority of the approaches described measure *relative* properties of biomolecules rather than *absolute* properties. Examples of this would include M_r estimation by mass spectrometry, gel filtration and electrophoresis, pI estimation by isoelectric focusing, secondary structure estimation of proteins by circular dichroism and determination of chemical shifts in NMR spectroscopy. For this reason, a common strategy found in many of the techniques is to *compare* the sample being analysed to a series of well-characterized standard molecules using well-established procedures which have been optimized for that particular method. It is important to understand that measurements obtained in this way are therefore highly dependent on standard measurements being of good quality and that this may vary somewhat from method to method.

1.5 THE WORLDWIDE WEB AS A RESOURCE IN PHYSICAL BIOCHEMISTRY

1.5.1 The Worldwide Web

The worldwide web was originally devised as a *distributed computer network* for the military capable of withstanding nuclear attack! In the last decade, it has grown to include millions of individual entries called *web pages* containing information on almost every conceivable subject. These web pages exist on a computer somewhere in the world but can in principle be accessed by other computers through the web. At the time of writing, (March, 2008) it is estimated that 20% (1320 million) of the Earth's 6606 million population have access to the web (http://www.internetworldstats.com/stats). We can connect to web pages on the web with an appropriate *browser* such as *Internet Explorer*. However, due to the sheer mass of material being constantly added to and changed on the web, we normally use a *search engine* to find pages on specific named topics. *Google* is a good example of a general purpose search engine. It should be remembered that no search engine gives 100% coverage so results from a search could represent as few as 25% of the total possible pages on a given topic.

The 'address' of a particular web page is given by a *uniform resource locator* (URL). Examples of URLs include http://www.google.com for google and http://alta-vista.com for altavista. The prefix http:// is to tell the receiving computer that it can expect a communication in *hypertext transfer protocol* – the most common format allowing one computer to communicate with another. The rest of the URL defines a location, that is a computer containing the relevant file. The ending .html which often occurs in URLs signifies *hypertext markup language*, the language in which web pages are written.

1.5.2 Web-Based Resources for Physical Biochemistry

The web provides several resources of use in Physical Biochemistry. Individual web pages are available which describe various experimental techniques thus complementing published work such as review articles and textbooks. There are also *databases* which are archives of one particular category of information. Examples would include sequence databases, databases of NMR spectra, the three dimensional structure database and databases of two-dimensional electrophoresis patterns. The best databases are *curated* (i.e. they are looked after and regularly updated by some reputable body) and they are *annotated* (which means each entry contains extra information such as literature citations, references to other related entries, etc.). These features make databases part of the daily life of modern molecular life scientists. Even though many resources on the web are not peer-reviewed in the way that say research articles are, most authoritative databases achieve the same result by maintaining a close link with the peer-reviewed literature. Conversely, it is becoming increasingly common for research articles to be submitted to journals in electronic format and for peer-reviewed articles to appear on the web long before the paper version. A third set of very useful resources on the web is made possible by the availability, often as *freeware* (i.e. for free!) of computer programs which help us to analyse or represent our data differently. Examples would include graphics programs which allow us to view the three dimensional structures of biomacromolecules encoded in *protein databank* (PDB) files (Chapters 6 and 9) or hydropathy plots which allow us to identify hydrophobic regions of amino acid sequences (Chapter 9).

The ever-closer links between molecular life sciences and *information technology* (IT) is represented in the relatively new discipline of *bioinformatics* which is introduced in Chapter 9. In this book relevant URLs for web-based resources are given at the end of each chapter.

In addition to text and programs, the wordwide web can be searched for images or videos using the standard search engines. A word of caution about using this type of searching in an academic context. The fact that material is on the web does not absolve us as scientists from respecting copyright law so permission should always be obtained to reproduce images, text or videos obtained from the web just as we would in using such content from a published source. Secondly, we should always take care to refer back to the primary literature as this is the bedrock of modern science and is likely to remain so as long as rigorous peer-review prevails.

1.6 OBJECTIVES OF THIS VOLUME

All of the techniques mentioned in this book merit one or more volumes to describe fully their potential for the future of life science research. In the bibliography at the end of each chapter the reader will find a list of such specialist texts and it is hoped that the present book will act as a general introduction to specialist biophysical techniques. In addition, recent review articles are cited which will bring the reader more up-to-date on specific applications of individual techniques. It is not the intention of the text to supplant such specialist literature but rather to guide students towards a greater understanding of the potential of biophysical approaches to biochemistry.

A chapter is devoted to each technique or group of techniques which describes the physical basis, advantages, limitations and opportunities it offers. This is presented in a generally nonmathematical way to maximize its accessibility (more detailed treatments may be found in the specialist texts). Moreover, the relationship between techniques is strongly emphasized because several combinations of individual techniques often offer advantages over single experimental approaches. In particular, recent advances have seen the combination of techniques such as mass spectrometry, chromatography, electrophoresis and spectroscopy as *hyphenated* or *multi-dimensional* analytical techniques. Care has also been taken to emphasize how biophysical approaches often complement biological and chemical experimentation to give a fuller understanding of biochemical systems.

Specific examples of applications of the approaches described are given in boxes throughout the text. These are meant to give a flavour of their versatility and power for the solving of many different types of problems in biochemistry. The bibliography contains many more examples such as, for example, applications in clinical laboratories and in industry. Articles and books (e.g. laboratory manuals) containing practical hints to novices contemplating using these techniques are also cited.

Finally, it is hoped that this book will furnish the student with sufficient understanding to allow them to understand and grasp as-yet undeveloped biophysical approaches which may appear in the next decade or so by noticing the common factors underlying the methods described as well as their diversity.

REFERENCES

Buffers and pH

Voit, E.O. and Ferreira, A.E.N. (1998) Buffering in models of integrated biochemical systems. *Journal of Theoretical Biology*, **191**, 429–37. A description of the effects of including buffers in modelling of biochemical systems.

Good, N.E. and Izawa, S. (1972) Hydrogen ion buffers for photosynthesis research. *Methods in Enzymology*, **XXIV**, 53–68. A description of Good's buffers.

Grady, J.K., Chasteen, N.D. and Harris, D.C. (1988) Radicals from Good's buffers. *Analytical Biochem.*, **173**, 111–5. Further reading on Good's buffers.

Units and quantities

Kotyk, A. (ed.) (1999) *Quantities, Symbols, Units and Abbreviations in the Life Sciences*. Humana Press, Totowa, NJ, USA. A guide to standard usage of units and quantities in the life sciences.

Methods for protein estimation

Bradford, M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical. Biochem.*, **72**, 248–54. Original description of the Bradford method.

Hartree, E.F. (1972) Determination of protein: A modification of the Lowry method that gives a alinear photometric response. *Analytical Biochem.*, **48**, 422–7. A later modification of the Lowry method.

Lowry, O., Rosebrough, A., Farr, A. and Randall, R. (1951) Protein measurement with the Folin phenol reagent. *J. Biol. Chem.*, **193**, 265–75. Original description of the Lowry method.

Williams, G.A., Macevilly, U., Ryan, R. and Harrington, M.G. (1995) Semiautomated protein assay using microtitre plates – some practical considerations. *British Journal of Biomedical Science*, **52**, 230–1. Description of problems encountered in miniaturising protein assays with microtitre plates.

Bioinformatics

Roberts, E., Eargle, J., Wright, D. and Luthey-Schulten, Z. (2006) MultiSeq: Unifying sequence and structure data for evolutionary analysis. *BMC Bioinformatics*, **7**, Art. No 382. This on-line journal article introduces MultiSeq, a tool for combining sequence and structure data for proteins including an excellent general discussion of current bioinformatic issues.

Lesk, A. (2005) *Introduction to Bioinformatics*, 2nd edn, Oxford University Press, Oxford, UK. An excellent, comprehensive and clear description of modern bioinformatics.

Some useful web sites

The SI system at National Institute of Standards and Technology (USA): http://www.physics.nist.gov/cuu/Units/.

An illustrated site in "chemguide" on ionization and acid-base chemistry by Jim Clark: http://www.chemguide.co.uk/physical/acideqiamenu.html.

Aquasol solubility database: http://www.pharm.arizona.edu/aquasol/index.html.

Useful bioinformatics sites: National Library of Medicine (USA): http://www.ncbi.nlm.nih.gov/.

European Bioinformatics Institute (UK): www.ebi.ac.uk.

Uniprot site (US/Europe): http://www.ebi.uniprot.org/index.shtml.

Chapter 2

Chromatography

Objectives

After completing this chapter you should be familiar with:

- *The* **physical basis** *of chromatography.*
- *The* **chemical basis** *of the principal chromatography methods used in biochemistry.*
- **Performance criteria** *which can be used to compare chromatography systems.*
- *The range of different* **chromatography formats** *used in biochemistry.*
- *How one might approach* **design of a purification protocol**, *for example to purify a specific protein of interest.*

Living cells contain hundreds of thousands of distinct chemical species. These include large molecules like proteins, nucleic acids, lipids as well as lower molecular weight molecules which act as building blocks for biopolymers or as components of complex metabolic pathways. Some of these molecules are present in only trace amounts (e.g. intermediates in enzyme mechanisms) whilst others are present in abundance (e.g. structural proteins). Moreover, some components are present only at certain stages of the cell-cycle, whilst others are present at approximately constant levels. Study of individual chemical components of cells can, therefore, give us an insight into many fundamental cellular processes and help us to understand the dynamics of cell composition and function.

One approach to the study of individual chemical species is to separate them from each other by analytical or preparative chromatography. Originally, this technique was used by Tswett (1903) in the separation of plant pigments (*chromatography* comes from the Greek, *chroma*, meaning *colour*) but we now know that it is applicable to all chemical species, whether coloured or not. Because of the large range of size, shape and hydrophobicity found in biomolecules, it is to be expected that no one chromatography technique will suffice for all separations. In this chapter, the basic principles of chromatography will be described to explain why different molecules are separable. Some examples of the main chromatographic techniques used in Biochemistry are then given to illustrate how biomolecules are separated in practice.

2.1 PRINCIPLES OF CHROMATOGRAPHY

2.1.1 The Partition Coefficient

When applied to any two-phase system (e.g. liquid–liquid, liquid–solid), a molecule may partition between the phases (Figure 2.1). The precise ratio of concentration achieved is ultimately determined by inherent thermodynamic properties of the molecule (in turn, a function of its chemical structure) and of the phases. In the case of a liquid–liquid system, the relative solubility of the molecule in each liquid will be very important in determining partitioning. In a liquid–solid system, different sample molecules may adsorb to varying degrees on the solid phase. Both partition and adsorption phenomena are possible in a column system and this is called column chromatography. In column chromatography, one phase is maintained stationary (the stationary phase) while the other (the mobile phase) may flow freely over it. We can express the concentration ratio in such a system as the partition coefficient, K:

$$K = \frac{C_s}{C_m} \tag{2.1}$$

where C_s and C_m are the sample concentrations in the stationary and mobile phases, respectively. When a mixture made up of several components is applied to such a two-phase system, each component will have its own individual

Physical Biochemistry: Principles and Applications, Second Edition David Sheehan

© 2009 John Wiley & Sons, Ltd

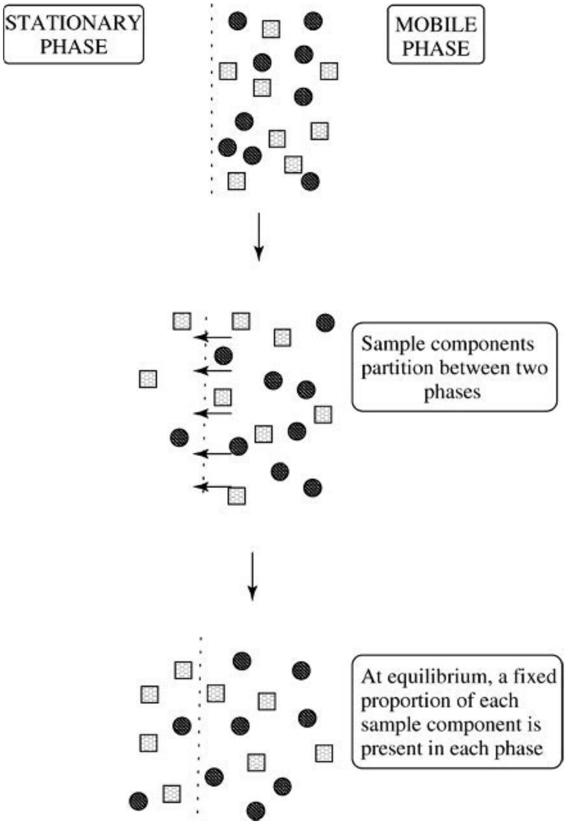
STATIONARY PHASE
MOBILE PHASE
Sample components partition between two phases
At equilibrium, a fixed proportion of each sample component is present in each phase

Figure 2.1. Partitioning of biomolecules in a two-phase system. Two components are represented by circles and squares, respectively. The two phases could be an aqueous buffer and a solid stationary phase. The two samples have very different partition coefficients in this experimental system.

partition coefficient. As a result, each will interact slightly differently with the stationary phase and, because of different partitioning between phases, will migrate through the column at different rates. Since K will be directly affected by the precise experimental conditions (e.g. temperature, solvent polarity) the chromatographer may vary these to optimize separation. In column chromatography, we therefore exploit what are often tiny differences in the partitioning behaviour of sample molecules to achieve their efficient separation.

2.1.2 Phase Systems Used in Biochemistry

In chromatographic systems used in biochemistry the stationary phase is made up of solid particles or of solid particles coated with liquid. In the former case, chemical groups are often covalently attached to the particles and this is called bonded phase liquid chromatography. In the latter case, a liquid phase may coat the particle and be attached by noncovalent, physical attraction. This type of system is called liquid–liquid phase chromatography. A good example of liquid–liquid phase chromatography is silica coated with a nonpolar hydrocarbon (e.g. C-18 reverse phase chromatography; Section 2.4.3). Commonly, the particles are composed of hydrated polymers such as cellulose or agarose. Such particles may be immobilized in a column (Section 2.3.1) and washed with mobile phase. They offer good flow characteristics and possess sufficient mechanical strength and chemical inertness for the chromatography of biomolecules. Because biomolecules have evolved to function in an aqueous environment, it is usually necessary to use aqueous buffers as the mobile phase if we require the molecule to retain its native structure (e.g. in the purification of active enzymes). If the native structure is not required, however, then it is possible to use more ‘nonbiological’ conditions such as organic solvents (e.g. in purification of peptides by reverse phase chromatography; Section 2.4.3).

Liquid–solid or liquid–liquid phases are the most common phase systems used in biochemistry. However, in specialized situations other phases may be used. For example, gas-solid and gas-liquid phases are used in gas chromatography (GC; Section 2.1.4). Regardless of the precise phase composition, chromatographic separation is a direct result of the different K values of each sample component.

2.1.3 Liquid Chromatography

To minimize loss of biological activity, separations are often carried out in aqueous buffers below room temperature. Low temperatures are especially important in the chromatography of cell extracts during, for example, protein purification. This reduces protease activity which might otherwise destroy the protein of interest. Chromatography with liquid mobile phases is called *liquid chromatography* (LC).

LC uses an experimental system outlined in Figure 2.2. Separation takes place in a column which contains the stationary phase. The volume and shape of the column will depend on the amount of sample to be separated and on the mode of chromatography to be used. Buffer is stored in a reservoir and is pumped through tubing onto the column. Appropriate valves allow the convenient injection of sample into this flow or the formation of gradients with a second buffer if required. The stationary phase is packed in the column and, as the sample passes through the bed of stationary phase, separation occurs. In *partition chromatography* modes, the sample separates into individual components as it passes through the stationary phase (e.g. *gel filtration*; Section 2.4.2). In *adsorption chromatography* modes, however, it is necessary first to load the entire sample and later to fractionate it. A good example of adsorption chromatography is *ion exchange chromatography* where the sample