

CONTENTS

Preface xiii

Abbreviations in Repetitive Use xvii

CHAPTER 1 INTRODUCTION 1

- 1.1 Prelude 1
- 1.2 Covalent Bonds 4
- 1.3 Noncovalent Interactions 5
 - 1.3.1 Electrostatic Interaction 6
 - 1.3.2 Van der Waals Interaction 6
 - 1.3.3 Hydrogen Bond 6
 - 1.3.4 Hydrophobic Interaction 7
 - 1.3.5 Steric Repulsion 8
- 1.4 Isomerism: Configuration versus Conformation 8
- 1.5 Trilogy 11
- 1.6 References 13

CHAPTER 2 MONOMER CONSTITUENTS OF BIOMACROMOLECULES 15

- 2.1 Nucleotides: Constituents of Nucleic Acids 15
- 2.2 α -Amino Acids: Constituents of Proteins 18
- 2.3 Monosaccharides: Constituents of Glycans 23
- 2.4 Addendum 28
- 2.5 References 30

CHAPTER 3 PURIFICATION AND CHARACTERIZATION 31

- 3.1 Purification: Overview 31
- 3.2 Purification: Chromatography 34
- 3.3 Purification: Electrophoresis 40
- 3.4 Characterization: General 44
 - 3.4.1 Purity 44
 - 3.4.2 Molecular Weight 44
 - 3.4.3 Molecular Dimension 50
- 3.5 Characterization: Specific 51
 - 3.5.1 Melting Temperature of DNA 51
 - 3.5.2 Buoyant Density of Biomacromolecules 52
 - 3.5.3 Isoelectric pH of Proteins 52
 - 3.5.4 Removal of Glycosides from Glycoproteins 53
- 3.6 References 53

CHAPTER 4 BIOMACROMOLECULAR STRUCTURE: NUCLEIC ACIDS 55

- 4.1 Structural Organization 55
 - 4.1.1 Structural Hierarchy 55
 - 4.1.2 Representation of Structures of Nucleic Acids 56
- 4.2 Sequence Analysis of Nucleic Acids 57
 - 4.2.1 General 57
 - 4.2.2 Chemical Cleavage Method 59
 - 4.2.3 Enzymatic Chain Termination/Dideoxy Method 60
 - 4.2.4 Mass Spectrometric Analysis 61
 - 4.2.5 Automated DNA Sequencing Technology 62
- 4.3 Secondary Structure and Structure Polymorphism of DNA 63
 - 4.3.1 Key Structural Features of Nucleic Acids 63
 - 4.3.2 DNA Polymorphism 66
 - 4.3.3 Alternative Structures of DNA 69
- 4.4 Supercoiling and Tertiary Structure of DNA 77
 - 4.4.1 DNA Topoisomers 77
 - 4.4.2 Superhelical Density and Energetics of Supercoiling 80
- 4.5 Classification and Structures of RNA 81
 - 4.5.1 Structures of RNA 81
 - 4.5.2 Transfer RNA 82
 - 4.5.3 Ribosomal RNA 83
 - 4.5.4 Messenger RNA 84
 - 4.5.5 Other Classes of RNA 85
- 4.6 RNA Folds and Structure Motifs 86
 - 4.6.1 RNA Folds 86
 - 4.6.2 Structure Motifs of RNA 86
- 4.7 Energetics of Nucleic Acid Structure 89
- 4.8 Nucleic Acid Application 90
- 4.9 References 91

CHAPTER 5 BIOMACROMOLECULAR STRUCTURE: PROTEINS 94

- 5.1 Architecture of Protein Molecules 94
 - 5.1.1 Introduction 94
 - 5.1.2 Representation of Protein Structures 94

5.2	Primary Structure of Proteins: Chemical and Enzymatic Sequence Analysis	95	5.9.1	Chromosomes	140
5.2.1	Amino Acid Composition	96	5.9.2	Ribosomes	141
5.2.2	Peptide Cleavage, Separation and Analysis	97	5.9.3	Spliceosome and Splicing Activities	142
5.2.3	Terminal and Sequence Determination	97	5.10	Conformational Energetics	143
5.2.4	Peptide Ladder Sequencing	101	5.11	References	144
5.3	Primary Structure of Proteins: Sequence Analysis by Tandem Mass Spectrometry	101	CHAPTER 6 BIOMACROMOLECULAR STRUCTURE: POLYSACCHARIDES 147		
5.3.1	An Application of Mass Spectrometry (MS) in Protein Chemistry	101	6.1	Propagation of Polysaccharide Chains	147
5.3.2	Application of Tandem Mass Spectrometry (MS-MS) in Protein Sequence Analysis	103	6.1.1	Introduction	147
5.4	Conformational Map	108	6.1.2	Representation of Glycan Structures	148
5.5	Secondary Structures and Motifs of Proteins	110	6.1.3	Toward Linear Code for Glycans	148
5.5.1	α -Helical Structure	111	6.2	Sequence Analysis of Polysaccharides: Primary Structure	153
5.5.2	β -Sheet Structure	113	6.2.1	Hydrolysis to Constituent Monosaccharides	154
5.5.3	Nonrepetitive Structure: Connection (Loop) and Turn	115	6.2.2	Chemical Methods	154
5.5.4	Notes to Secondary Structures of Globular Proteins	116	6.2.3	Enzymatic Methods	155
5.5.5	Motifs: Supersecondary Structures	117	6.2.4	Spectrometric Methods	157
5.6	Domains and Tertiary Structures of Proteins	118	6.3	Conformation: Secondary and Tertiary Structures of Polysaccharide Chains	161
5.6.1	Domain Structures	119	6.4	Conformation: Description of Some Polysaccharide Structures	163
5.6.2	Tertiary Structures and Protein Folds	121	6.4.1	Starch	163
5.6.3	Folds and Protein Binding	126	6.4.2	Glycogen	164
5.6.4	Membrane Proteins	128	6.4.3	Pectins	165
5.6.5	Fibrous Proteins	128	6.4.4	Cellulose	165
5.6.6	Circular (Cyclic) Proteins	129	6.4.5	Chitin	166
5.6.7	Representation of Protein Topology	130	6.5	Glycobiology: Study of Glycoprotein-Associated Glycans	167
5.6.8	Accessible Surface of Folded Structures	130	6.5.1	Glycoprotein and Glycoforms	167
5.7	Classification of Protein Structures	133	6.5.2	Structure Diversity of Oligosaccharide Chains	168
5.7.1	α -Helical Proteins	133	6.5.3	Structural Analysis	174
5.7.2	β -Sheet Proteins	133	6.6	Neoglycoproteins	177
5.7.3	$\alpha + \beta$ Proteins	135	6.7	Organizational Levels of Biomacromolecular Structures	177
5.7.4	α/β Proteins	135	6.8	References	181
5.7.5	Multidomain Structures	135	CHAPTER 7 STUDIES OF BIOMACROMOLECULAR STRUCTURES: SPECTROSCOPIC ANALYSIS OF CONFORMATION 183		
5.7.6	Membrane and Cell Surface Proteins	136	7.1	Biochemical Spectroscopy: Overview	183
5.7.7	Irregular and Small Proteins	136	7.2	Ultraviolet and Visible Absorption Spectroscopy	185
5.8	Quaternary (Subunit) Structures of Proteins	137	7.2.1	Basic Principles	185
5.9	Quinternary Structure Exemplified: Nucleoproteins	140	7.2.2	Amino Acid Residues and Peptide Bonds	187
			7.2.3	Purines, Pyrimidines and Nucleic Acids	188

7.2.4	Perturbation Difference Absorption Spectroscopy	189
7.3	Fluorescence Spectroscopy	190
7.4	Infrared Spectroscopy	193
7.4.1	Basic Principles	193
7.4.2	Biochemical Applications	195
7.5	Nuclear Magnetic Resonance Spectroscopy	197
7.5.1	Basic Principles	197
7.5.2	Two-Dimensional Fourier Transform NMR	202
7.5.3	NMR of Proteins	203
7.5.4	NMR of Nucleic Acids	206
7.5.5	NMR of Glycans	207
7.6	Optical Rotatory Dispersion and Circular Dichroism Spectroscopy	208
7.6.1	Basic Principles	208
7.6.2	ORD/CD Spectra and Protein Secondary Structures	209
7.6.3	Empirical Applications of ORD and CD	212
7.7	X-ray Diffraction Spectroscopy	214
7.7.1	Basic Principles	214
7.7.2	Crystallographic Study of Biomacromolecules	216
7.8	References	219

CHAPTER 8 STUDIES OF BIOMACROMOLECULAR STRUCTURES: CHEMICAL SYNTHESIS 220

8.1	Rationale	220
8.2	Synthetic Strategy: Conventional Approach	220
8.2.1	Protection and Deprotection of Common Functional Groups	221
8.2.2	Protection and Deprotection Specific to Peptide Synthesis	223
8.2.3	Coupling Reaction	225
8.3	Synthetic Strategy: Solid Phase Approach	225
8.3.1	General Concept	225
8.3.2	Solid-Phase Polymer Support	230
8.4	Practice of Solid Phase Synthesis and Its Application	232
8.4.1	Oligo- and Polypeptide Synthesis	232
8.4.2	Oligo- and Polynucleotide Synthesis	236
8.4.3	Oligo- and Polysaccharide Synthesis	237
8.5	Combinatorial Synthesis	241
8.5.1	Parallel Synthesis	241
8.5.2	Mixture Synthesis	242

8.6	Biochemical Polypeptide Chain Ligation	245
8.7	References	247

CHAPTER 9 STUDIES OF BIOMACROMOLECULAR STRUCTURES: COMPUTATION AND MODELING 249

9.1	Potential Energy and Molecular Thermodynamics	249
9.2	Molecular Modeling: Molecular Mechanical Approach	252
9.2.1	Introduction	252
9.2.2	Energy Calculation	254
9.2.3	Energy Minimization	256
9.2.4	Molecular Dynamics	258
9.2.5	Conformational Search	261
9.2.6	Remaining Issues	262
9.2.7	Computational Application of Molecular Modeling Packages	263
9.3	Statistical Thermodynamics	264
9.3.1	General Principles	264
9.3.2	Transitions of Regular Structures: Two-State Models	268
9.3.3	Random Structure: Random-Walk Problem	271
9.4	Structural Transition: Examples	273
9.4.1	Coil-Helix Transition in Polypeptides	273
9.4.2	Helical Transition in Nucleic Acids	274
9.4.3	Topological Transition of Closed Circular DNA Duplex	276
9.5	Structure Prediction from Sequence by Statistical Methods	276
9.5.1	Approaches	276
9.5.2	Secondary Structure of Proteins and Beyond	277
9.5.3	Functional Sites of Proteins	280
9.5.4	Nucleic Acid Fold	281
9.6	Molecular Docking: Prediction of Biomacromolecular Binding	282
9.7	References	286

CHAPTER 10 BIOMACROMOLECULAR INTERACTION 289

10.1	Biomacromolecules in Solution	289
10.2	Multiple Equilibria	291
10.2.1	Single-Site Binding	291
10.2.2	Multiple-Site Binding: General	292
10.2.3	Multiple-Site Binding: Equivalent Sites	293

10.2.4	Multiple-Site Binding: Nonequivalent Sites	294
10.3	Allosterism and Cooperativity	295
10.3.1	Models	295
10.3.2	Diagnostic Tests for Cooperativity	299
10.4	Specificity and Diversity of Antibody- Antigen Interactions	300
10.4.1	Structure of Antibody	300
10.4.2	Antibody-Antigen Complex	303
10.5	Complementarity in Nucleic Acid Interactions	305
10.5.1	DNA-Protein Interaction	305
10.5.2	Binding of Intercalation Agent to Supercoiled DNA	309
10.5.3	RNA-Protein Interaction	310
10.6	Molecular Recognition in Carbohydrate- Lectin Interaction	312
10.6.1	Classification and Structures of Lectins	312
10.6.2	Lectin-Carbohydrate Recognition: General	315
10.6.3	Lectin-Carbohydrate Recognition: Ligand Discrimination	318
10.7	References	320

CHAPTER 11 BIOMACROMOLECULAR CATALYSIS 322

11.1	Biocatalyst: Definition and Classification	322
11.2	Characteristics of Enzymes	325
11.2.1	Enzymes: Catalytic Proteins	325
11.2.2	Catalytic Efficiency	326
11.2.3	Enzyme Specificity	328
11.2.4	Active Site of Enzyme	330
11.2.5	Multienzyme Complex and Multifunctional Enzymes	331
11.3	Enzyme Kinetics	333
11.3.1	Fundamental of Enzyme Kinetics	333
11.3.2	Steady-State Kinetic Treatment of Enzyme Catalysis	336
11.3.3	Quasi-Equilibrium Treatment of Random Reactions	338
11.3.4	Cleland's Approach	339
11.3.5	Nonlinear Kinetics	339
11.3.6	Environmental Effects	341
11.4	Enzyme Mechanisms	344
11.4.1	Essay on Enzyme Reaction Mechanism	344
11.4.2	Studies of Enzyme Mechanism: Active Site	349
11.4.3	Studies of Enzyme Mechanism: Transition State	356

11.4.4	Structure-Activity Relationship	357
11.4.5	X-ray Crystallographic Studies and Refinement	361
11.4.6	Case Studies of Enzyme Mechanisms	361
11.5	Enzyme Regulation	374
11.5.1	Elements of Enzyme Regulation	374
11.5.2	Covalent Modifications of Enzymes and Cascade Effect	374
11.5.3	Control of Enzyme Catalytic Activity by Effectors	377
11.5.4	Structure Basis of Allosteric Regulation: Glycogen Phosphorylase	381
11.6	Abzyme	383
11.7	Ribozyme	386
11.7.1	Characteristics of Catalytic RNA	386
11.7.2	Description of Ribozymes	388
11.7.3	Strategies for Ribozyme Catalysis	392
11.8	References	394

CHAPTER 12 SIGNAL TRANSDUCTION AND BIODEGRADATION 398

12.1	Chemical Transduction: Metabolism	398
12.2	Elements of Signal Transduction	400
12.2.1	First Messengers	400
12.2.2	Receptors	400
12.2.3	Second Messengers	403
12.2.4	Transducers: GTP-Binding Proteins	403
12.3	Effector Enzymes and Signal Transduction	406
12.3.1	Adenylyl Cyclase and Signal Transduction	406
12.3.2	Phospholipase C and Signal Transduction	408
12.4	Topics on Signal Transduction	410
12.4.1	Calcium Signaling	410
12.4.2	Phosphorylation and Dephosphorylation in Signaling	414
12.4.3	Signal Pathways Operated by Receptor Protein Tyrosine Kinase	417
12.4.4	Signaling Pathways Operated by Nonreceptor Proteins Tyrosine Kinase	419
12.5	Apoptosis	419
12.6	Hydrolysis versus Phosphorolysis of Glycans	422

12.7	Nucleolysis of Nucleic Acids	424
12.8	Proteolysis and Protein Degradation	426
12.8.1	Proteolytic Mechanism	426
12.8.2	Protein Degradation Pathway	427
12.9	References	433

CHAPTER 13 BIOSYNTHESIS AND GENETIC TRANSMISSION 436

13.1	Saccharide Biosynthesis and Glycobiology	436
13.1.1	Biosynthesis of Biopolymer: Distributive versus Processive	436
13.1.2	Biosynthesis of oligo- and polysaccharide chains	436
13.1.3	Biosynthesis of Glycoproteins	437
13.2	Genetic Information and Transmission	442
13.3	DNA Replication and Repair	445
13.3.1	DNA Replication: Overview	445
13.3.2	DNA Replication: Enzymology	448
13.3.3	Reverse Transcription	455
13.3.4	Post-Replicational Modification	456
13.3.5	DNA Repair	458
13.4	Biosynthesis and Transcription of RNA	461
13.4.1	RNA Transcription: Prokaryotic System	461
13.4.2	RNA Transcription: Eukaryotic System	463
13.4.3	Regulation of RNA Transcription	466
13.4.4	Posttranscriptional Processing/Modification	469
13.5	Translation and Protein Biosynthesis	472
13.5.1	Protein Translation: Overview	472
13.5.2	Protein Translation: Processes	475
13.5.3	Decoding Mechanism	479
13.5.4	Recoding, Frameshifting and Expanded Genetic Code	481
13.5.5	Rescue System for Stalled Ribosomes	483
13.5.6	Posttranslational Modifications of Protein	484
13.5.7	Protein Translocation	488
13.6	Folding of Biomacromolecules	491
13.6.1	Overview	491
13.6.2	RNA Folding	491
13.6.3	<i>In vitro</i> Protein Folding Pathway	492
13.6.4	Molecular Chaperone in Cytosolic Protein Folding	494

13.7	Bioengineering of Biomacromolecules	494
13.7.1	Recombinant DNA Technology	494
13.7.2	RNA Engineering	500
13.7.3	Protein Engineering	501
13.7.4	Antibody Engineering	506
13.8	References	511

CHAPTER 14 BIOMACROMOLECULAR INFORMATICS 515

14.1	Overview	515
14.2	Biosequences	515
14.2.1	Sequencing Biomacromolecules	515
14.2.2	Sequence Similarity and Pair-Wise Alignment	517
14.2.3	Similarity Search and Multiple Sequence Alignment	522
14.2.4	Statistical Significance of Sequence Search/Alignments	524
14.3	Microarray: General Description	525
14.3.1	Introduction	525
14.3.2	Surface Preparation for Microarray	525
14.3.3	Microarray Targets	528
14.3.4	Microarray Probes	529
14.3.5	Biochemical Reaction of Microarray	530
14.3.6	Microarray Detection	530
14.3.7	Data analysis in microarray	531
14.4	Computer Technology	533
14.4.1	Machine: Computer	533
14.4.2	Tool: Program, Language and Programming	535
14.4.3	Molecular Graphics	537
14.4.4	Resource: Internet	540
14.3.5	Internet Resources of Biochemical Interest	546
14.5	Informatics	548
14.5.1	Introduction to Database	548
14.5.2	Biochemical Databases	549
14.5.3	Database Retrieval	551
14.6	Gene Ontology	553
14.7	References	555

CHAPTER 15 GENOMICS 558

15.1	Genome: Features and Organization	558
15.1.1	Genome Features	558
15.1.2	Gene Mapping	561
15.1.3	Information Content of Nucleotide Sequence	563
15.1.4	DNA Library	564
15.1.5	Alternative Splicing	566

15.1.6	Gene Variation: Single Nucleotide Polymorphism	567
15.2	Genome Informatics: Databases and Web Servers	568
15.2.1	Nucleic Acid Databases	568
15.2.2	Nucleic Acid Analysis Servers	571
15.3	Approaches to Gene Identification	571
15.3.1	Masking Repetitive DNA	575
15.3.2	Database Searches	576
15.3.3	Codon Bias Detection	576
15.3.4	Detecting Functional Sites in the DNA	577
15.4	Gene Expression	578
15.4.1	Expression Profiling: DNA Chips	578
15.4.2	Gene Expression: mRNA Quantification and Transcriptome Analysis	583
15.5	Genome Project	587
15.6	References	590

CHAPTER 16 PROTEOMICS 594

16.1	Proteome: Features and Properties	594
16.1.1	Proteome Features	594
16.1.2	Protein Identity Based on Composition and Properties	595
16.1.3	Physicochemical Properties Based on Sequence	596
16.2	Proteome Informatics: Sequence Databases and Servers	598
16.2.1	Amino Acid Sequence	598
16.2.2	Primary Sequence Database	599
16.2.3	Secondary Sequence Database	602
16.2.4	Boutique Databases	605
16.3	Proteome Informatics: Structure Databases and Servers	605
16.3.1	Structure Database: Primary Archive	605
16.3.2	Structure Databases: Substructures and Structure Classification	608
16.4	Proteome Informatics: Proteomic Servers	610
16.4.1	Proteome Analysis and Annotation	610
16.4.2	Integrated Databases	613
16.4.3	Post-Translational Modifications and Functional Sites	614
16.5	Protein Structure Analysis Using Bioinformatics	616
16.5.1	Secondary Structure Predictions	617

16.5.2	Three-Dimensional Structure Modeling	618
16.5.3	Sequence Similarity and Alignment	619
16.5.4	Structure Similarity and Overlap	620
16.5.5	Fold Recognition and Threading	623
16.5.6	Homology Modeling	623
16.5.7	<i>Ab initio</i> Prediction of Protein Structure	624
16.5.8	Solvation	625
16.5.9	On-line Protein Structure Prediction	626
16.5.10	Protein-Protein Interaction	628
16.6	Investigation of Proteome Expression and Function	629
16.6.1	Two-Dimensional Gel Electrophoresis	629
16.6.2	Proteome Analysis by Mass Spectrometry	631
16.6.3	Analysis of Posttranslational Modification by Mass Spectrometry	634
16.6.4	High Throughput Protein Crystallography	635
16.6.5	Protein-Protein Interactions by Two-Hybrid Assay	636
16.6.6	Protein Chip	638
16.6.7	Activity-Based Probe	640
16.6.8	Nonsense Suppression Mutagenesis	643
16.7	Metabolome	647
16.8	References	650

CHAPTER 17 GLYCOMICS 655

17.1	Features of Glycomics	655
17.1.1	Glycobiology: Nomenclature and Representation of Glycans	655
17.1.2	Glycobiology: Glycoforms	657
17.1.3	Glycomics: Response to Post-Genomic Era	659
17.2	Glycomic Databases and Servers	661
17.2.1	Glycan Structure	661
17.2.2	Glycan Analysis	663
17.2.3	Glycosylation of Proteins	665
17.3	Glycomics: Genetic Approaches	666
17.4	Glycomics: Proteoglycomic Approaches	668
17.4.1	Characterization of Glycosylation Sites	668
17.4.2	Lectin and Glycoenzyme-Based Proteoglycomics	670

17.4.3	Metabolic Oligosaccharide Engineering	672	18.3.5	Construction of Phylogenetic Tree	692
17.4.4	Recombinant Glycoproteins	673	18.3.6	Assessment	692
17.5	Glycomics: Chemoglycomic Approaches	674	18.4	Application of Sequence Analyses in Phylogenetic Inference	693
17.5.1	Structural Analysis of Glycans	674	18.4.1	Phylogenetic Analysis Software	693
17.5.2	Glycoprotein Syntheses in Glycomics	674	18.4.2	Phylogenetic Analysis with PHYLIP	693
17.5.3	Glycochip	675	18.4.3	Phylogenetic Analysis Online	697
17.6	References	678	18.5	Evolution of Biosequences	697
CHAPTER 18 BIOMACROMOLECULAR EVOLUTION			18.5.1	Evolution of Nucleic Acid Sequence	697
			18.5.2	Regulation of Evolutionary Change	699
18.1	Variation in Biomacromolecular Sequences	680	18.6	Evolution of Protein Structure and Function	701
18.1.1	Mutation as Driving Force of Evolution	680	18.6.1	Evolution of Protein Complexity: General	702
18.1.2	Evolutionary Rate and Role of Selection	682	18.6.2	Evolution of Protein Complexity: Domain Duplication	703
18.2	Element of Molecular Phylogeny	685	18.6.3	Evolution of Protein Structure: Fold Change	704
18.3	Phylogenetic Analysis of Biosequences	687	18.6.4	Evolution of Protein Function: Catalytic Site Convergence versus Divergence	706
18.3.1	General Consideration	687	18.7	References	708
18.3.2	Sequence Data	687	INDEX		710
18.3.3	Phylogenetic Method: Distance-Based Approaches	690			
18.3.4	Phylogenetic Method: Character-Based Approaches	691			