

CONTENTS

Preface	xix
Common Abbreviations	xxi
Glossary	xxv
Section I: Macromolecules	
1. Configurations and conformations	1
1.1. Stereochemistry	2
1.1.A. Chirality	3
1. Enantiomers	5
2. Racemic mixtures	6
3. Diastereomers	7
4. Epimers and epimerization	8
5. <i>Cis</i> and <i>trans</i> isomers	9
1.1.B. Prochiral	10
1.1.C. Tautomers	12
1.2. Conformations	13
1.2.A. Torsion angle	15
1.2.B. Dihedral angle	16
1.3. Conformations of idealized polymers	17
1.3.A. Random coils	18
1. End-to-end distances	19
2. Radius of gyration	21
3. Characteristic ratio	21
1.3.B. Excluded volume effects and theta solvents	22
1. Covalent cross-links	23
1.4. Structure databases: structures on the WEB	24
Section II: Nucleic acids	
2. DNA structure	25
2.1. Polynucleotides	26

2.1.A. The deoxyribose group	29
2.1.B. Properties of the bases	31
2.1.C. Modifications of the bases	37
2.2. DNA three-dimensional structures	40
2.2.A. Base pairing and stacking	42
2.2.B. Double helices	46
1. B-DNA	52
2. A-DNA	53
3. Z-DNA	54
2.2.C. Other DNA structures	56
1. Hoogsteen base pairs	56
2. Triple helices	57
3. H-DNA: intramolecular triple helices	60
4. Four-stranded structures: guanine quartet	60
5. i- motif	63
6. Inverted repeat sequences and palindromes	64
7. Helical junctions: cruciforms, Holliday junctions	66
8. Parallel-stranded DNA duplexes	66
2.3. DNA as a polyelectrolyte: hydration and counterions	68
2.4. DNA flexibility and dynamics: curving, twisting, stretching	71
2.4.A. Local flexibility	74
2.4.B. Hydrogen exchange	75
2.5. Binding of small molecules	77
2.5.A. Binding to the minor groove	77
2.5.B. Intercalation	78
1. Ethidium bromide	78
2. Psoralen photo cross-linking	79
2.6. Chemical modification as a probe of structure	80
2.6.A. Cross-linking	85
3. DNA topology	86
3.1. Supercoiling and superhelices: topoisomers	90
3.2. Linking number, Lk	92
3.2.A. Linking difference, ΔLk	92
1. Relaxed duplex DNA	93
3.2.B. Superhelix density, σ	94
3.3. Topoisomerases	95
3.4. Twist and writhe	96
3.4.A. Twist, Tw	96
3.4.B. Writhe, Wr	97
3.5. DNA topology and geometry	98
3.5.A. Experimental characterization of DNA topology	99
1. Electron microscopy	99
2. Gel electrophoresis to separate topoisomers	101

3. Intercalation by ethidium bromide	101
4. Two-dimensional gel electrophoresis	101
3.6. Energetics of supercoiling	103
3.6. B. Energy distribution of topoisomers	104
3.6. C. Topology-dependent binding of ligands	105
3.7. DNA wrapped around the nucleosome	106
4. RNA structure	108
4.1. Secondary structure of RNA	112
4.1.A. Hairpin loops	116
4.1.B. Tetraloops	117
4.1.C. Bulges and internal loops	118
4.2. Tertiary structure of RNA	118
4.2.A. Common structural motifs	120
1. Pseudoknots	120
2. Coaxial helices: interhelical stacking	121
3. A-minor motif	122
4. Dinucleotide platform	123
5. “Kissing” hairpin loops	123
6. Ribose zipper	124
7. Uridine turn	125
8. Tetraloop/receptor interactions	126
9. Roles of ions	126
4.2.B. Transfer RNA structures	127
4.2.C. Ribozyme structures	129
4.3. Quaternary structure of RNA	133
4.4. RNA structure prediction	134
4.4.A. Prediction of secondary structure	134
1. Thermodynamic approach	134
2. Phylogenetic approach	135
4.4.B. Prediction of tertiary structure	136
5. Denaturation, renaturation, and hybridization of nucleic acids	139
5.1. Denaturation of double-stranded nucleic acids	139
5.1.A. Methods for monitoring denaturation	140
5.1.B. Double-stranded DNA	142
1. Thermal melting	143
2. Denaturants	144
3. pH	145
4. Salt effects	145
5. Prediction of the T_m	146
5.1.C. Double-stranded RNA	148
5.1.D. DNA • RNA heteroduplexes	149
5.1.E. Single-stranded nucleic acids	149
1. Physical stretching	150
5.2. Unfolding and refolding of single-stranded RNA molecules	150

5.2.A. Transfer RNA unfolding/refolding	153
5.2.B. Ribozyme unfolding/refolding	154
5.2.C. Unfolding using mechanical force	155
5.3. Renaturation, annealing, and hybridization	157
5.3.A. Competing intramolecular structures in individual single strands	162
5.3.B. C_0t and R_0t curves	163
5.3.C. Probe hybridization	165
1. Stringency	167
2. Analyzing the extent of complementarity	168
3. <i>In situ</i> hybridization	170
5.4. DNA mimics: peptide nucleic acids	170
5.4.A. Chemistry and synthesis	172
5.4.B. Hybridization properties	172
1. PNA•DNA and PNA•RNA duplexes	173
2. $(PNA)_2$ •DNA triplexes	173
3. Strand invasion: binding to double-stranded DNA	173
4. PNA duplexes and triplexes	175
5.4.C. Structures of PNA complexes	175
6. Manipulating nucleic acids	177
6.1. Replicating DNA	177
6.1.A. DNA polymerase	178
6.1.B. DNA ligase	181
6.1.C. Polymerase chain reaction (PCR)	182
6.2. Producing RNA	185
6.2.A. RNA replication: RNA replicases	185
6.2.B. Transcription: DNA-dependent RNA polymerases	186
1. Single-subunit phage DNA-dependent RNA polymerases	188
6.2.C. Reverse transcription: RNA into DNA	189
6.2.D. Antisense oligonucleotides	190
6.3. Cloning	191
6.3.A. Expression vectors	193
6.3.B. cDNA libraries	194
6.3.C. Restriction enzymes	195
6.3.D. Restriction maps	196
6.4. Sequencing DNA	198
6.4.A. Isolating the DNA fragments to be sequenced	199
6.4.B. Chain-termination, Sanger method	199
6.4.C. Separating the DNA fragments by size	203
6.4.D. Alternative approaches	204
6.5. Sequencing RNA	205
6.5.A. Direct sequencing of oligoribonucleotides	205
6.5.B. Identifying modified nucleotides	207
6.6. Chemical synthesis of DNA	208
6.6.A. Protecting groups for 2'-deoxynucleosides	211
6.6.B. Coupling methods	211

1. Phosphotriester procedure	211
2. Phosphoramidite procedure	212
3. <i>H</i> -Phosphonate procedure	213
6.6.C. Solution-phase DNA synthesis	215
6.6.D. Solid-phase DNA synthesis	216
6.6.E. Site-directed mutagenesis	218
6.7. Chemical synthesis of RNA	220
6.7.A. Protecting the 2'-hydroxyl group	220
6.7.B. RNA synthesis in solution	222
6.7.C. Solid-phase RNA synthesis	223

Section III: Proteins

7. Polypeptide structure	227
7.1. Polypeptide chains	227
7.2. Amino acid residues	229
7.2.A. Glycine (Gly)	230
7.2.B. Nonpolar amino acid residues (Ala, Leu, Ile, Val)	232
7.2.C. Hydroxyl residues (Ser, Thr)	232
7.2.D. Arginine (Arg)	233
7.2.E. Lysine (Lys)	234
1. Acetylation by anhydrides	236
2. Amidination	236
3. Guanidination	237
4. Schiff base formation	238
5. Carbamylation	239
7.2.F. Histidine (His)	239
7.2.G. Acidic residues (Asp, Glu)	241
7.2.H. Amide residues (Asn, Gln)	242
1. Deamidation	243
7.2.I. Cysteine (Cys)	245
1. Alkylation of thiol groups	245
2. Thiol addition across double bonds	246
3. Binding of metal ions	247
4. Oxidation of thiol groups	247
5. Disulfide bonds	248
6. Thiol-disulfide exchange	249
7. Dithiothreitol, dithioerythritol	252
8. Ellman's reagent	253
7.2.J. Methionine (Met)	254
7.2.K. Phenylalanine (Phe)	255
7.2.L. Tyrosine (Tyr)	256
7.2.M. Tryptophan (Trp)	257
7.2.N. Imino acid (Pro)	258
7.2.O. Selenocysteine (Sec)	259
7.2.P. Physical properties and hydrophobicities of amino acid residues	260

1. Hydrophilicities	261
2. Hydrophobicities	261
7.3. Protein detection	267
7.3.A. Biuret reaction	267
7.3.B. Lowry assay	268
7.3.C. Ninhydrin	268
7.3.D. Fluorescamine	269
7.3.E. Coomassie brilliant blue	270
7.3.F. Ponceau S	271
7.4. Peptide synthesis	271
7.4.A. Chemistry of polypeptide chain assembly	272
1. Chemical ligation of peptide fragments	275
7.4.B. Solution or solid phase?	276
7.4.C. Peptide libraries	278
7.5. Peptide and protein sequencing	280
7.5.A. Amino acid analysis	280
1. Peptide bond hydrolysis	281
2. Quantifying amino acids	282
3. Counting residues	283
7.5.B. Fragmentation of a protein into peptides	285
1. Proteolytic enzymes	285
2. Chemical methods of cleavage	288
7.5.C. Peptide mapping	289
7.5.D. Diagonal maps	291
1. Isolating peptides containing certain amino acids	291
2. Identifying disulfide bonds	292
7.5.E. Sequencing	293
1. Amino-terminal and carboxyl-terminal residues	294
2. Sequencing from the N-terminus: the Edman degradation	295
3. Sequencing from the C-terminus	298
4. Sequencing by mass spectrometry	299
7.5.F. Protein sequences from gene sequences	301
1. Post-translational modifications	302
7.6. Primary structures of natural proteins: evolution at the molecular level	306
7.6.A. Homologous genes and proteins.	307
1. Detecting sequence homology	310
2. Aligning homologous sequences	313
3. Orthologous / paralogous genes and proteins	314
4. Nature of amino acid sequence differences	315
5. Rates of divergence	318
6. Roles of selection.	321
a. Neutral mutations and negative selection	322
b. Positive selection for functional mutations	323
7.6.B. Gene rearrangements and the evolution of protein complexity.	324
1. Gene duplications	324

2. Protein elongation by intra-gene duplication	325
3. Gene fusion and division	326
7.6.C. Protein engineering	326
8. Polypeptide conformation	329
8.1. Local flexibility of the polypeptide backbone: the Ramachandran plot	329
8.2. Random-coil polypeptide chains	334
8.2.A. Statistical properties	334
8.2.B. Rates of conformational change	336
1. X-Pro peptide bond <i>cis/trans</i> isomerization	337
8.3. Regular structures	338
8.3.A. α -Helix	339
1. 3_{10} - and Π -helices	342
8.3.B. β -Sheet	342
8.3.C. Polyglycine	344
8.3.D. Polyproline	344
8.4 α -Helix formation from a random coil	346
8.4.A. Factors stabilizing the α -helix	347
8.4.B. Helix-coil transitions	348
8.4.C. Helix-coil models	351
1. Zimm-Bragg model	351
2. Lifson-Roig model	352
8.4.D. Trifluoroethanol	353
8.5. Fibrous proteins	353
8.5.A α -Fibrous structures: coiled coils	354
8.5.B. β -Fibrous structures	358
8.5.C. Collagens	359
9. Protein structure	362
9.1. Three-dimensional structures of globular proteins: molecular complexity	363
9.1.A. Tertiary structure: the overall fold	366
9.1.B. Secondary structure: regular local structures	369
1. Helices	371
2. β -Structure	372
3. β -Bulge	374
9.1.C. Reverse turns: changing direction	375
1. β -Turns.	376
2. γ -Turns.	377
3. Omega loops	378
9.1.D. Supersecondary structures: common motifs	378
1. β -Hairpin and β -meander	379
2. β -helix and β -roll: β -solenoids	380
3. Cystine knot	381
4. Greek key motif	381
5. Jelly roll motif	382
6. Four-helix bundle	383

7. Epidermal growth factor (EGF) motif	384
8. Interleukin-1 motif: trefoils	384
9. Kringle domain	385
9.1.E. Contact Maps	386
9.1.F. Interiors and exteriors.	388
9.1.G. The solvent: interactions with water	393
9.1.H. Quaternary structure: initiating macromolecular assembly	394
1. Symmetry	398
2. Asymmetry and approximate symmetry	400
3. Interfaces	402
4. Oligomerization and domain swapping	403
5. Filamentous arrays	405
9.1.I. Classification of tertiary structures: order out of diversity	405
1. α Structures	405
2. β Structures	406
3. $\alpha\beta$ Structures	407
a. TIM, $(\alpha/\beta)_8$ barrels	407
4. $\alpha + \beta$ and other structures	408
5. Protein structure classification databases	408
9.2. Membrane proteins: avoiding water	409
9.2.A. Helical superfolds	411
9.2.B. β -Barrel membrane proteins	413
9.2.C. Monotopic and bitopic membrane proteins	414
9.2.D. Interactions with the membrane	414
9.3. Proteins with similar folded conformations: evolution in 3-D	415
9.3.A. Homologous proteins: protein families	415
1. Structural homology within a polypeptide chain	421
9.3.B. Structural similarity without apparent sequence homology: surprises	421
9.3.C. Sequence similarity without structural homology: new folds?	422
9.4. Protein structure prediction	424
9.4.A. <i>Ab initio</i> predictions: the ultimate goal	425
9.4.B. Secondary structure prediction: a one-dimensional problem	426
1. Identifying transmembrane helices: hydrophathy	427
9.4.C. Homology modeling	429
1. Inverse folding problem	430
2. Threading protein sequences	430
9.4.D. <i>De novo</i> protein design	432
10. Physical properties of folded proteins	434
10.1. Solubilities and volumes of proteins in water	436
10.1.A. Hydration layer	437
10.1.B. Partial volumes	438
10.2. Chemical reactivities	441
10.2.A. Ionization: electrostatic effects	445
10.3. Isotope (hydrogen) exchange	447

10.3.A. Exchange in macromolecules	447
10.3.B. Solvent penetration model	449
10.3.C. Local unfolding mechanism	449
1. EX1 mechanism	450
2. EX2 mechanism	450
10.4. Flexibility detected crystallographically	451
10.4.A. Effects of different crystal lattices	451
10.4.B. The temperature factor: mobility or disorder?	452
10.5. Flexibility detected by NMR	453
10.5.A. NMR time scales	453
10.5.B. Aromatic ring flipping	454
10.6. Varying the temperature	456
10.7. Effects of high pressure	457
10.7.A. Adiabatic compressibility	457
10.7.B. Isothermal compressibility	458
10.7.C. Structural effects of high pressure	458
10.8. Spectral properties	459
10.9. Integral membrane proteins	460
11. Protein denaturation: unfolding and refolding	462
11.1. Reversible unfolding at equilibrium	463
11.1.A. Reversibility of denaturation	463
11.1.B. Cooperativity of unfolding	465
11.1.C. Denaturants	468
11.1.D. Heat denaturation	472
11.1.E. Cold denaturation	474
11.1.F. pH denaturation	476
11.1.G. Denaturation by high pressure	478
11.1.H. Breakage of disulfide bonds	480
11.2. Unfolded proteins	480
11.2.A. Molten globule	484
11.2.B. Conformational equilibria in polypeptide fragments	486
11.3. Protein stability	488
11.3.A. Physical basis of protein stability	491
1. Water and co-solvents	496
a. Stabilizers	496
b. Destabilizers	497
c. The role of water	497
11.3.B. Effects of varying the primary structure	498
1. Natural proteins of exceptional stability	498
2. Mutagenic studies	499
11.3.C. Structural stability of membrane proteins	503
11.4. Protein refolding <i>in vitro</i>	504
11.4.A. Refolding of single-domain proteins	505
1. Characterizing the transition state for folding	507
2. Kinetic schemes for folding	509

11.4.B. Kinetic determination of folding	510
1. Bacterial proteinases	511
2. Serpin proteinase inhibitors	512
11.4.C. Folding coupled to disulfide formation	513
11.4.D. Proteins with multiple domains	516
11.4.E. Proteins with multiple subunits	516
11.4.F. Competition with aggregation and precipitation	517
11.4.G. Differences with folding <i>in vivo</i>	518
Section IV: Functions	
12. Ligand binding by proteins	519
12.1. General properties of protein-ligand interactions	520
12.2. Metalloproteins	525
12.2.A. Chelation: synergy between ligands	527
12.2.B. Zinc-binding proteins	529
12.2.C. Metallothioneins	531
12.2.D. Iron-transport and storage proteins	532
12.2.E. Blue-copper proteins	534
12.3. Calcium-binding proteins	535
12.3.A. EF-hand calcium-binding proteins	536
1. Calmodulin and troponin C	538
12.3.B. Carboxylation and hydroxylation of Asp, Asn and Glu residues	539
12.4. NAD- and nucleotide-binding proteins	540
12.4.A. Dinucleotide binding motif	541
12.4.B. Mononucleotide-binding motif	543
12.5. Allostery: interactions between different binding sites	543
12.5.A. Structural models	544
1. Sequential model: direct interactions	544
2. Concerted model: quaternary structure changes	545
3. Comparison of the sequential and concerted models	546
12.5.B. Hemoglobin and myoglobin	547
1. Structure	548
2. Oxygen binding	549
3. Cooperativity of oxygen binding	552
4. Heterotropic interactions	554
5. Bohr effect	555
6. Allosteric mechanism of hemoglobin	557
12.5.C. Negative cooperativity	560
1. Negative cooperativity or heterogeneity of sites?	562
13. Nucleic acid/protein interactions	564
13.1. Techniques for measuring protein-DNA interactions	566
13.1.A. Filter-binding assays	567
13.1.B. Gel retardation assay	568
13.1.C. Footprinting	570
13.2. Principles of protein-DNA recognition	572

13.2.A. Specificity of DNA-protein binding	573
1. Specific interactions	576
2. Nonspecific complexes	578
3. Water-mediated contacts	579
4. Dehydration effects	579
5. Release of condensed counterions	581
13.2.B. Changes in the protein conformation	582
13.2.C. Changes in the DNA conformation	582
13.3. DNA-binding structural motifs	585
13.3.A. Helix-turn-helix motif	587
1. Lac repressor	590
2. Lambda CI and Cro repressors	591
3. Homeodomains	592
4. POU domains	593
5. Trp repressor	595
6. Cyclic AMP receptor protein (CRP) / Catabolite gene activator protein (CAP)	596
13.3.B. TATA-binding protein (TBP)	598
13.3.C. Zinc-containing DNA-binding motifs	599
1. Zinc fingers	599
2. Steroid hormone receptors	602
3. GAL4 type	603
13.3.D. bZip and helix-loop-helix domains	604
13.3.E. β -Sheets: Methionine repressor	605
13.3.F. Histone fold	607
13.3.G. Bacterial type-II DNA-binding proteins: heat-unstable (HU) and integration host factor (IHF)	608
13.3.H. Single-strand DNA-binding proteins	609
1. Prokaryotic single-strand DNA binding proteins	609
2. Eukaryotic replication protein A	610
3. OB (oligonucleotide/oligosaccharide binding) fold	611
13.4. RNA-binding proteins	613
13.4.A. Ribonucleoprotein (RNP) domain	616
13.4.B. Double-stranded RNA-binding domain	616
13.4.C. KH domain	618
13.4.D. MS2 bacteriophage coat protein	619
13.4.E. Recognizing transfer RNAs	620
1. Class-I glutamyl-tRNA synthetase	621
2. Class-II aspartyl-tRNA synthetase	622
3. Class-II seryl-tRNA synthetase	623
4. Elongation factor EF-Tu	624
13.4.F. The ribosome	625
14. Catalysis	628
14.1. Chemical catalysis	630

14.2. Enzyme kinetics: Michaelis-Menten	630
14.2.A. The Michaelis-Menten equation	633
14.2.B. K_m (Michaelis constant)	635
14.2.C. Turnover number (k_{cat})	635
14.2.D. Lineweaver-Burke plot	635
1. Eadie-Hofstee plot	636
2. Lineweaver-Burke <i>versus</i> Eadie-Hofstee	637
14.2.E. Kinetics of individual enzyme molecules	637
14.3. Enzyme kinetic mechanisms with multiple substrates	638
14.3.A. Sequential mechanisms	639
1. Ordered mechanisms	640
2. Random mechanisms	640
14.3.B. Non-sequential mechanisms: Ping-Pong	640
14.3.C. Initial rate equations	641
1. Steady-state ordered and rapid-equilibrium mechanisms	641
2. Equilibrium ordered mechanism	643
3. Ping-pong mechanism	643
14.3.D. Dead-end inhibitors	645
14.3.E. Competitive inhibition	646
1. Linear competitive inhibition	647
2. Hyperbolic competitive inhibition	649
14.3.F. Noncompetitive inhibition	649
1. Linear noncompetitive inhibition	649
2. Hyperbolic noncompetitive inhibition	650
14.3.G. Uncompetitive inhibition	651
14.3.H. Substrate inhibition	652
14.3.I. Product inhibition	654
14.3.J. Haldane relationship	655
14.3.K. Isotope exchange at equilibrium	656
1. Ping-pong mechanism	657
2. Sequential reactions	658
14.3.L. Slow- and tight-binding enzyme inhibitors	659
14.4. Mechanisms of enzyme catalysis	662
14.4.A. Reactions on the enzyme	666
14.4.B. Stabilizing the transition state	668
1. Transition state analogues	669
14.4.C. Entropic contributions	672
14.4.D. Bisubstrate analogues	675
1. Ap_5A	675
2. <i>N</i> -phosphonacetyl-L-aspartate (PALA)	676
14.4.E. Induced fit	677
14.4.F. Covalent catalysis	680
14.4.G. Cofactors, coenzymes, and prosthetic groups	681
1. Pyridoxal phosphate	681
14.4.H. Suicide substrates	683

14.4.I. Cryoenzymology	685
14.4.J. Time-resolved crystallography	685
14.4.K. Polymeric substrates: processivity	686
14.4.L. Enzyme function <i>in vivo</i> : toward “perfection”	688
14.4.M. One example: tyrosyl tRNA synthetase	691
1. Editing of amino acid activation	695
14.5. Catalytic antibodies	696
14.6. Catalytic nucleic acids: ribozymes and deoxyribozymes (DNAzymes).	698
14.6.A. Natural reactions	698
14.6.B. Ribozyme structure and catalysis	701
14.6.C. Selection for novel ribozymes and deoxyribozymes	702
14.6.D. Ligand-binding nucleic acids: aptamers	704
15. Enzyme regulation	706
15.1. Allosteric enzymes	707
15.1.A. Allosteric models	708
15.1.B. Structural aspects	709
15.1.C. Aspartate transcarbamoylase	710
15.1.D. Phosphofructokinase	717
1. Mechanism of phosphoryl transfer	718
2. Allosteric properties of <i>E. coli</i> PFK	719
15.1.E. Threonine synthase	720
15.2. Covalent regulation	721
15.2.A. Phosphorylation	721
1. Phosphorylation in eukaryotes	722
2. Phosphorylation in prokaryotes	724
3. Protein phosphorylation in signal transduction networks	725
4. Specificity of protein phosphorylation	725
5. Effects of phosphorylation on the properties of proteins	726
6. Methods to characterize protein phosphorylation	727
15.2.B. Glycogen phosphorylase	728
1. Regulation by phosphorylation	731
2. Allosteric properties of glycogen phosphorylase	733
15.2.C. Adenylylation	735
1. Glutamine synthetase	736
15.2.D. Proteolysis: turning zymogens into proteinases	738
1. Trypsin family of serine proteinases	738
2. Carboxyl proteinases	740
3. Metalloproteinases	742