

Preface of the First Edition *xv*
Preface of the Second Edition *xvii*

1	Networks in Biological Cells	1
1.1	Some Basics About Networks	1
1.1.1	Random Networks	2
1.1.2	Small-World Phenomenon	2
1.1.3	Scale-Free Networks	3
1.2	Biological Background	4
1.2.1	Transcriptional Regulation	5
1.2.2	Cellular Components	5
1.2.3	Spatial Organization of Eukaryotic Cells into Compartments	7
1.2.4	Considered Organisms	8
1.3	Cellular Pathways	8
1.3.1	Biochemical Pathways	8
1.3.2	Enzymatic Reactions	11
1.3.3	Signal Transduction	11
1.3.4	Cell Cycle	12
1.4	Ontologies and Databases	12
1.4.1	Ontologies	12
1.4.2	Gene Ontology	13
1.4.3	Kyoto Encyclopedia of Genes and Genomes	13
1.4.4	Reactome	13
1.4.5	Brenda	14
1.4.6	DAVID	14
1.4.7	Protein Data Bank	15
1.4.8	Systems Biology Markup Language	15
1.5	Methods for Cellular Modeling	17
1.6	Summary	17
1.7	Problems	17
	Bibliography	18

2	Structures of Protein Complexes and Subcellular Structures	21
2.1	Examples of Protein Complexes	22
2.1.1	Principles of Protein–Protein Interactions	24
2.1.2	Categories of Protein Complexes	27
2.2	Complexome: The Ensemble of Protein Complexes	28
2.2.1	Complexome of <i>Saccharomyces cerevisiae</i>	28
2.2.2	Bacterial Protein Complexomes	30
2.2.3	Complexome of Human	31
2.3	Experimental Determination of Three-Dimensional Structures of Protein Complexes	31
2.3.1	X-ray Crystallography	32
2.3.2	NMR	34
2.3.3	Electron Crystallography/Electron Microscopy	34
2.3.4	Cryo-EM	34
2.3.5	Immunoelectron Microscopy	35
2.3.6	Fluorescence Resonance Energy Transfer	35
2.3.7	Mass Spectroscopy	36
2.4	Density Fitting	38
2.4.1	Correlation-Based Density Fitting	38
2.5	Fourier Transformation	40
2.5.1	Fourier Series	40
2.5.2	Continuous Fourier Transform	41
2.5.3	Discrete Fourier Transform	41
2.5.4	Convolution Theorem	41
2.5.5	Fast Fourier Transformation	42
2.6	Advanced Density Fitting	44
2.6.1	Laplacian Filter	45
2.7	FFT Protein–Protein Docking	46
2.8	Protein–Protein Docking Using Geometric Hashing	48
2.9	Prediction of Assemblies from Pairwise Docking	49
2.9.1	CombDock	49
2.9.2	Multi-LZerD	52
2.9.3	3D-MOSAIC	52
2.10	Electron Tomography	53
2.10.1	Reconstruction of Phantom Cell	55
2.10.2	Protein Complexes in <i>Mycoplasma pneumoniae</i>	55
2.11	Summary	56
2.12	Problems	57
2.12.1	Mapping of Crystal Structures into EM Maps	57
	Bibliography	60
3	Analysis of Protein–Protein Binding	63
3.1	Modeling by Homology	63
3.2	Properties of Protein–Protein Interfaces	66
3.2.1	Size and Shape	66
3.2.2	Composition of Binding Interfaces	68

3.2.3	Hot Spots	69
3.2.4	Physicochemical Properties of Protein Interfaces	71
3.2.5	Predicting Binding Affinities of Protein–Protein Complexes	72
3.2.6	Forces Important for Biomolecular Association	73
3.3	Predicting Protein–Protein Interactions	75
3.3.1	Pairing Propensities	75
3.3.2	Statistical Potentials for Amino Acid Pairs	78
3.3.3	Conservation at Protein Interfaces	79
3.3.4	Correlated Mutations at Protein Interfaces	83
3.4	Summary	86
3.5	Problems	86
	Bibliography	86
4	Algorithms on Mathematical Graphs	89
4.1	Primer on Mathematical Graphs	89
4.2	A Few Words About Algorithms and Computer Programs	90
4.2.1	Implementation of Algorithms	91
4.2.2	Classes of Algorithms	92
4.3	Data Structures for Graphs	93
4.4	Dijkstra’s Algorithm	95
4.4.1	Description of the Algorithm	96
4.4.2	Pseudocode	100
4.4.3	Running Time	101
4.5	Minimum Spanning Tree	101
4.5.1	Kruskal’s Algorithm	102
4.6	Graph Drawing	102
4.7	Summary	104
4.8	Problems	105
4.8.1	Force Directed Layout of Graphs	107
	Bibliography	110
5	Protein–Protein Interaction Networks – Pairwise Connectivity	111
5.1	Experimental High-Throughput Methods for Detecting Protein–Protein Interactions	111
5.1.1	Gel Electrophoresis	112
5.1.2	Two-Dimensional Gel Electrophoresis	112
5.1.3	Affinity Chromatography	113
5.1.4	Yeast Two-hybrid Screening	114
5.1.5	Synthetic Lethality	115
5.1.6	Gene Coexpression	116
5.1.7	Databases for Interaction Networks	116
5.1.8	Overlap of Interactions	116
5.1.9	Criteria to Judge the Reliability of Interaction Data	118
5.2	Bioinformatic Prediction of Protein–Protein Interactions	120
5.2.1	Analysis of Gene Order	121
5.2.2	Phylogenetic Profiling/Coevolutionary Profiling	121

5.2.2.1	Coevolution	122
5.3	Bayesian Networks for Judging the Accuracy of Interactions	124
5.3.1	Bayes' Theorem	125
5.3.2	Bayesian Network	125
5.3.3	Application of Bayesian Networks to Protein–Protein Interaction Data	126
5.3.3.1	Measurement of Reliability “Likelihood Ratio”	127
5.3.3.2	Prior and Posterior Odds	127
5.3.3.3	A Worked Example: Parameters of the Naïve Bayesian Network for Essentiality	128
5.3.3.4	Fully Connected Experimental Network	129
5.4	Protein Interaction Networks	131
5.4.1	Protein Interaction Network of <i>Saccharomyces cerevisiae</i>	131
5.4.2	Protein Interaction Network of <i>Escherichia coli</i>	131
5.4.3	Protein Interaction Network of Human	132
5.5	Protein Domain Networks	132
5.6	Summary	135
5.7	Problems	136
5.7.1	Bayesian Analysis of (Fake) Protein Complexes	136
	Bibliography	138

6 Protein–Protein Interaction Networks – Structural Hierarchies 141

6.1	Protein Interaction Graph Networks	141
6.1.1	Degree Distribution	141
6.1.2	Clustering Coefficient	143
6.2	Finding Cliques	145
6.3	Random Graphs	146
6.4	Scale-Free Graphs	147
6.5	Detecting Communities in Networks	149
6.5.1	Divisive Algorithms for Mapping onto Tree	153
6.6	Modular Decomposition	155
6.6.1	Modular Decomposition of Graphs	157
6.7	Identification of Protein Complexes	161
6.7.1	MCODE	161
6.7.2	ClusterONE	162
6.7.3	DACO	163
6.7.4	Analysis of Target Gene Coexpression	164
6.8	Network Growth Mechanisms	165
6.9	Summary	169
6.10	Problems	169
	Bibliography	178

7 Protein–DNA Interactions 181

7.1	Transcription Factors	181
7.2	Transcription Factor-Binding Sites	183

7.3	Experimental Detection of TFBS	183
7.3.1	Electrophoretic Mobility Shift Assay	183
7.3.2	DNAse Footprinting	184
7.3.3	Protein-Binding Microarrays	185
7.3.4	Chromatin Immunoprecipitation Assays	187
7.4	Position-Specific Scoring Matrices	187
7.5	Binding Free Energy Models	189
7.6	<i>Cis</i> -Regulatory Motifs	191
7.6.1	DACO Algorithm	192
7.7	Relating Gene Expression to Binding of Transcription Factors	192
7.8	Summary	194
7.9	Problems	194
	Bibliography	195

8	Gene Expression and Protein Synthesis	197
8.1	Regulation of Gene Transcription at Promoters	197
8.2	Experimental Analysis of Gene Expression	198
8.2.1	Real-time Polymerase Chain Reaction	199
8.2.2	Microarray Analysis	199
8.2.3	RNA-seq	201
8.3	Statistics Primer	201
8.3.1	<i>t</i> -Test	203
8.3.2	<i>z</i> -Score	203
8.3.3	Fisher's Exact Test	203
8.3.4	Mann–Whitney–Wilcoxon Rank Sum Tests	205
8.3.5	Kolmogorov–Smirnov Test	206
8.3.6	Hypergeometric Test	206
8.3.7	Multiple Testing Correction	207
8.4	Preprocessing of Data	207
8.4.1	Removal of Outlier Genes	207
8.4.2	Quantile Normalization	208
8.4.3	Log Transformation	208
8.5	Differential Expression Analysis	209
8.5.1	Volcano Plot	210
8.5.2	SAM Analysis of Microarray Data	210
8.5.3	Differential Expression Analysis of RNA-seq Data	212
8.5.3.1	Negative Binomial Distribution	213
8.5.3.2	DESeq	213
8.6	Gene Ontology	214
8.6.1	Functional Enrichment	216
8.7	Similarity of GO Terms	217
8.8	Translation of Proteins	217
8.8.1	Transcription and Translation Dynamics	218
8.9	Summary	219
8.10	Problems	220
	Bibliography	224

9	Gene Regulatory Networks	227
9.1	Gene Regulatory Networks (GRNs)	228
9.1.1	Gene Regulatory Network of <i>E. coli</i>	228
9.1.2	Gene Regulatory Network of <i>S. cerevisiae</i>	231
9.2	Graph Theoretical Models	231
9.2.1	Coexpression Networks	232
9.2.2	Bayesian Networks	233
9.3	Dynamic Models	234
9.3.1	Boolean Networks	234
9.3.2	Reverse Engineering Boolean Networks	235
9.3.3	Differential Equations Models	236
9.4	DREAM: Dialogue on Reverse Engineering Assessment and Methods	238
9.4.1	Input Function	239
9.4.2	YAYG Approach in DREAM3 Contest	240
9.5	Regulatory Motifs	244
9.5.1	Feed-forward Loop (FFL)	245
9.5.2	SIM	245
9.5.3	Densely Overlapping Region (DOR)	246
9.6	Algorithms on Gene Regulatory Networks	247
9.6.1	Key-pathway Miner Algorithm	247
9.6.2	Identifying Sets of Dominating Nodes	248
9.6.3	Minimum Dominating Set	249
9.6.4	Minimum Connected Dominating Set	249
9.7	Summary	250
9.8	Problems	251
	Bibliography	254
10	Regulatory Noncoding RNA	257
10.1	Introduction to RNAs	257
10.2	Elements of RNA Interference: siRNAs and miRNAs	259
10.3	miRNA Targets	261
10.4	Predicting miRNA Targets	264
10.5	Role of TFs and miRNAs in Gene-Regulatory Networks	264
10.6	Constructing TF/miRNA Coregulatory Networks	266
10.6.1	TFmiR Web Service	267
10.6.1.1	Construction of Candidate TF–miRNA–Gene FFLs	268
10.6.1.2	Case Study	269
10.7	Summary	270
	Bibliography	270
11	Computational Epigenetics	273
11.1	Epigenetic Modifications	273
11.1.1	DNA Methylation	273
11.1.1.1	CpG Islands	276
11.1.2	Histone Marks	277
11.1.3	Chromatin-Regulating Enzymes	278

11.1.4	Measuring DNA Methylation Levels and Histone Marks Experimentally	279
11.2	Working with Epigenetic Data	281
11.2.1	Processing of DNA Methylation Data	281
11.2.1.1	Imputation of Missing Values	281
11.2.1.2	Smoothing of DNA Methylation Data	281
11.2.2	Differential Methylation Analysis	282
11.2.3	Comethylation Analysis	283
11.2.4	Working with Data on Histone Marks	285
11.3	Chromatin States	286
11.3.1	Measuring Chromatin States	286
11.3.2	Connecting Epigenetic Marks and Gene Expression by Linear Models	287
11.3.3	Markov Models and Hidden Markov Models	288
11.3.4	Architecture of a Hidden Markov Model	290
11.3.5	Elements of an HMM	291
11.4	The Role of Epigenetics in Cellular Differentiation and Reprogramming	292
11.4.1	Short History of Stem Cell Research	293
11.4.2	Developmental Gene Regulatory Networks	293
11.5	The Role of Epigenetics in Cancer and Complex Diseases	295
11.6	Summary	296
11.7	Problems	296
	Bibliography	301

12 Metabolic Networks 303

12.1	Introduction	303
12.2	Resources on Metabolic Network Representations	306
12.3	Stoichiometric Matrix	308
12.4	Linear Algebra Primer	309
12.4.1	Matrices: Definitions and Notations	309
12.4.2	Adding, Subtracting, and Multiplying Matrices	310
12.4.3	Linear Transformations, Ranks, and Transpose	311
12.4.4	Square Matrices and Matrix Inversion	311
12.4.5	Eigenvalues of Matrices	312
12.4.6	Systems of Linear Equations	313
12.5	Flux Balance Analysis	314
12.5.1	Gene Knockouts: MOMA Algorithm	316
12.5.2	OptKnock Algorithm	318
12.6	Double Description Method	319
12.7	Extreme Pathways and Elementary Modes	324
12.7.1	Steps of the Extreme Pathway Algorithm	324
12.7.2	Analysis of Extreme Pathways	328
12.7.3	Elementary Flux Modes	329
12.7.4	Pruning Metabolic Networks: NetworkReducer	331
12.8	Minimal Cut Sets	332
12.8.1	Applications of Minimal Cut Sets	337

12.9	High-Flux Backbone	339
12.10	Summary	341
12.11	Problems	341
12.11.1	Static Network Properties: Pathways	341
	Bibliography	346
13	Kinetic Modeling of Cellular Processes	349
13.1	Biological Oscillators	349
13.2	Circadian Clocks	350
13.2.1	Role of Post-transcriptional Modifications	352
13.3	Ordinary Differential Equation Models	353
13.3.1	Examples for ODEs	354
13.4	Modeling Cellular Feedback Loops by ODEs	356
13.4.1	Protein Synthesis and Degradation: Linear Response	356
13.4.2	Phosphorylation/Dephosphorylation – Hyperbolic Response	357
13.4.3	Phosphorylation/Dephosphorylation – Buzzer	359
13.4.4	Perfect Adaptation – Sniffer	360
13.4.5	Positive Feedback – One-Way Switch	361
13.4.6	Mutual Inhibition – Toggle Switch	362
13.4.7	Negative Feedback – Homeostasis	362
13.4.8	Negative Feedback: Oscillatory Response	364
13.4.9	Cell Cycle Control System	365
13.5	Partial Differential Equations	366
13.5.1	Spatial Gradients of Signaling Activities	368
13.5.2	Reaction–Diffusion Systems	368
13.6	Dynamic Phosphorylation of Proteins	369
13.7	Summary	370
13.8	Problems	372
	Bibliography	373
14	Stochastic Processes in Biological Cells	375
14.1	Stochastic Processes	375
14.1.1	Binomial Distribution	376
14.1.2	Poisson Process	377
14.1.3	Master Equation	377
14.2	Dynamic Monte Carlo (Gillespie Algorithm)	378
14.2.1	Basic Outline of the Gillespie Method	379
14.3	Stochastic Effects in Gene Transcription	380
14.3.1	Expression of a Single Gene	380
14.3.2	Toggle Switch	381
14.4	Stochastic Modeling of a Small Molecular Network	385
14.4.1	Model System: Bacterial Photosynthesis	385
14.4.2	Pools-and-Proteins Model	386
14.4.3	Evaluating the Binding and Unbinding Kinetics	387
14.4.4	Pools of the Chromatophore Vesicle	389
14.4.5	Steady-State Regimes of the Vesicle	389
14.5	Parameter Optimization with Genetic Algorithm	392

14.6	Protein–Protein Association	395
14.7	Brownian Dynamics Simulations	396
14.8	Summary	398
14.9	Problems	400
14.9.1	Dynamic Simulations of Networks	400
	Bibliography	407
15	Integrated Cellular Networks	409
15.1	Response of Gene Regulatory Network to Outside Stimuli	410
15.2	Whole-Cell Model of <i>Mycoplasma genitalium</i>	412
15.3	Architecture of the Nuclear Pore Complex	416
15.4	Integrative Differential Gene Regulatory Network for Breast Cancer Identified Putative Cancer Driver Genes	416
15.5	Particle Simulations	421
15.6	Summary	423
	Bibliography	424
16	Outlook	427
	Index	429