Contents

Preface xi
Guide to Different Topics of the Book xiii
About the Authors xv

Part One Introduction to Systems Biology 1

1 Introduction 3
  1.1 Biology in Time and Space 3
  1.2 Models and Modeling 4
    1.2.1 What Is a Model? 4
    1.2.2 Purpose and Adequateness of Models 5
    1.2.3 Advantages of Computational Modeling 5
  1.3 Basic Notions for Computational Models 6
    1.3.1 Model Scope 6
    1.3.2 Model Statements 6
    1.3.3 System State 6
    1.3.4 Variables, Parameters, and Constants 6
    1.3.5 Model Behavior 7
    1.3.6 Model Classification 7
    1.3.7 Steady States 7
    1.3.8 Model Assignment Is Not Unique 7
  1.4 Networks 8
  1.5 Data Integration 8
  1.6 Standards 9
  1.7 Model Organisms 9
    1.7.1 Escherichia coli 9
    1.7.2 Saccharomyces cerevisiae 11
    1.7.3 Caenorhabditis elegans 11
    1.7.4 Drosophila melanogaster 11
    1.7.5 Mus musculus 12
    References 12
    Further Reading 14

2 Modeling of Biochemical Systems 15
  2.1 Overview of Common Modeling Approaches for Biochemical Systems 15
  2.2 ODE Systems for Biochemical Networks 17
    2.2.1 Basic Components of ODE Models 18

2.2.2 Illustrative Examples of ODE Models 18
    References 21
    Further Reading 21

3 Structural Modeling and Analysis of Biochemical Networks 23
  3.1 Structural Analysis of Biochemical Systems 24
    3.1.1 System Equations 24
    3.1.2 Information Encoded in the Stoichiometric Matrix N 25
    3.1.3 The Flux Cone 27
    3.1.4 Elementary Flux Modes and Extreme Pathways 27
    3.1.5 Conservation Relations – Null Space of $N^T$ 29
  3.2 Constraint-Based Flux Optimization 30
    3.2.1 Flux Balance Analysis 31
    3.2.2 Geometric Interpretation of Flux Balance Analysis 31
    3.2.3 Thermodynamic Constraints 31
    3.2.4 Applications and Tests of the Flux Optimization Paradigm 32
    3.2.5 Extensions of Flux Balance Analysis 33
    Exercises 35
    References 36
    Further Reading 37

4 Kinetic Models of Biochemical Networks: Introduction 39
  4.1 Reaction Kinetics and Thermodynamics 39
    4.1.1 Kinetic Modeling of Enzymatic Reactions 39
    4.1.2 The Law of Mass Action 40
    4.1.3 Reaction Thermodynamics 40
    4.1.4 Michaelis–Menten Kinetics 42
    4.1.5 Regulation of Enzyme Activity by Effectors 44
    4.1.6 Generalized Mass Action Kinetics 48
    4.1.7 Approximate Kinetic Formats 48
    4.1.8 Convenience Kinetics and Modular Rate Laws 49
  4.2 Metabolic Control Analysis 50
    4.2.1 The Coefficients of Control Analysis 51
## Contents

4.2.2 The Theorems of Metabolic Control Theory 53
4.2.3 Matrix Expressions for Control Coefficients 55
4.2.4 Upper Glycolysis as Realistic Model Example 58
4.2.5 Time-Dependent Response Coefficients 59
Exercises 61
References 61
Further Reading 62

5 Data Formats, Simulation Techniques, and Modeling Tools 63
5.1 Simulation Techniques and Tools 63
5.1.1 Differential Equations 63
5.1.2 Stochastic Simulations 64
5.1.3 Simulation Tools 65
5.2 Standards and Formats for Systems Biology 72
5.2.1 Systems Biology Markup Language 72
5.2.2 BioPAX 74
5.2.3 Systems Biology Graphical Notation 74
5.3 Data Resources for Modeling of Cellular Reaction Systems 75
5.3.1 General-Purpose Databases 75
5.3.2 Pathway Databases 76
5.3.3 Model Databases 77
5.4 Sustainable Modeling and Model Semantics 78
5.4.1 Standards for Systems Biology Models 78
5.4.2 Model Semantics and Model Comparison 78
5.4.3 Model Combination 80
5.4.4 Model Validity 82
References 83
Further Reading 85

6 Model Fitting, Reduction, and Coupling 87
6.1 Parameter Estimation 88
6.1.1 Regression, Estimators, and Maximal Likelihood 88
6.1.2 Parameter Identifiability 90
6.1.3 Bootstrapping 91
6.1.4 Bayesian Parameter Estimation 92
6.1.5 Probability Distributions for Rate Constants 94
6.1.6 Optimization Methods 97
6.2 Model Selection 99
6.2.1 What Is a Good Model? 99
6.2.2 The Problem of Model Selection 100
6.2.3 Likelihood Ratio Test 102
6.2.4 Selection Criteria 102
6.2.5 Bayesian Model Selection 103
6.3 Model Reduction 104
6.3.1 Model Simplification 104
6.3.2 Reduction of Fast Processes 105
6.3.3 Quasi-Equilibrium and Quasi-Steady State 107
6.3.4 Global Model Reduction 108

6.4 Coupled Systems and Emergent Behavior 110
6.4.1 Modeling of Coupled Systems 111
6.4.2 Combining Rate Laws into Models 113
6.4.3 Modular Response Analysis 113
6.4.4 Emergent Behavior in Coupled Systems 114
6.4.5 Causal Interactions and Global Behavior 115
Exercises 116
References 117
Further Reading 119

7 Discrete, Stochastic, and Spatial Models 121
7.1 Discrete Models 122
7.1.1 Boolean Networks 122
7.1.2 Petri Nets 124
7.2 Stochastic Modeling of Biochemical Reactions 127
7.2.1 Chance in Biochemical Reaction Systems 127
7.2.2 The Chemical Master Equation 129
7.2.3 Stochastic Simulation 129
7.2.4 Chemical Langevin Equation and Chemical Noise 130
7.2.5 Dynamic Fluctuations 132
7.2.6 From Stochastic to Deterministic Modeling 133
7.3 Spatial Models 133
7.3.1 Types of Spatial Models 134
7.3.2 Compartment Models 135
7.3.3 Reaction–Diffusion Systems 136
7.3.4 Robust Pattern Formation in Embryonic Development 138
7.3.5 Spontaneous Pattern Formation 139
7.3.6 Linear Stability Analysis of the Activator–Inhibitor Model 140
Exercises 142
References 143
Further Reading 144

8 Network Structure, Dynamics, and Function 145
8.1 Structure of Biochemical Networks 146
8.1.1 Random Graphs 147
8.1.2 Scale-Free Networks 148
8.1.3 Connectivity and Node Distances 149
8.1.4 Network Motifs and Significance Tests 150
8.1.5 Explanations for Network Structures 151
8.2 Regulation Networks and Network Motifs 152
8.2.1 Structure of Transcription Networks 153
8.2.2 Regulation Edges and Their Steady-State Response 156
8.2.3 Negative Feedback 156
8.2.4 Adaptation Motif 157
8.2.5 Feed-Forward Loops 158
8.3 Modularity and Gene Functions 160
8.3.1 Cell Functions Are Reflected in Structure, Dynamics, Regulation, and Genetics 160
8.3.2 Metabolism Pathways and Elementary Modes 162
8.3.3 Epistasis Can Indicate Functional Modules 163
8.3.4 Evolution of Function and Modules 163
8.3.5 Independent Systems as a Tacit Model Assumption 165
8.3.6 Modularity and Biological Function Are Conceptual Abstractions 165
Exercises 166
References 167
Further Reading 169

9 Gene Expression Models 171
9.1 Mechanisms of Gene Expression Regulation 171
9.1.1 Transcription Factor-Initiated Gene Regulation 171
9.1.2 General Promoter Structure 173
9.1.3 Prediction and Analysis of Promoter Elements 174
9.1.4 Posttranscriptional Regulation through microRNAs 176
9.2 Dynamic Models of Gene Regulation 180
9.2.1 A Basic Model of Gene Expression and Regulation 180
9.2.2 Natural and Synthetic Gene Regulatory Networks 183
9.2.3 Gene Expression Modeling with Stochastic Equations 186
9.3 Gene Regulation Functions 187
9.3.1 The Lac Operon in E. coli 187
9.3.2 Gene Regulation Functions Derived from Equilibrium Binding 188
9.3.3 Thermodynamic Models of Promoter Occupancy 189
9.3.4 Gene Regulation Function of the Lac Promoter 191
9.3.5 Inferring Transcription Factor Activities from Transcription Data 192
9.3.6 Network Component Analysis 194
9.3.7 Correspondences between mRNA and Protein Levels 196
9.4 Fluctuations in Gene Expression 196
9.4.1 Stochastic Model of Transcription and Translation 197
9.4.2 Intrinsic and Extrinsic Variability 200
9.4.3 Temporal Fluctuations in Gene Cascades 202
Exercises 203
References 205
Further Reading 207

10 Variability, Robustness, and Information 209
10.1 Variability and Biochemical Models 210
10.1.1 Variability and Uncertainty Analysis 210
10.1.2 Flux Sampling 212
10.1.3 Elasticity Sampling 213
10.1.4 Propagation of Parameter Variability in Kinetic Models 214
10.1.5 Models with Parameter Fluctuations 216
10.2 Robustness Mechanisms and Scaling Laws 217
10.2.1 Robustness in Biochemical Systems 218
10.2.2 Robustness by Backup Elements 219
10.2.3 Feedback Control 219
10.2.4 Perfect Robustness by Structure 222
10.2.5 Scaling Laws 224
10.2.6 Time Scaling, Summation Laws, and Robustness 227
10.2.7 The Role of Robustness in Evolution and Modeling 228
10.3 Adaptation and Exploration Strategies 229
10.3.1 Information Transmission in Signaling Pathways 230
10.3.2 Adaptation and Fold-Change Detection 230
10.3.3 Two Adaptation Mechanisms: Sensing and Random Switching 231
10.3.4 Shannon Information and the Value of Information 232
10.3.5 Metabolic Shifts and Anticipation 233
10.3.6 Exploration Strategies 234
Exercises 236
References 237
Further Reading 239

11 Optimality and Evolution 241
11.1 Optimality in Systems Biology Models 243
11.1.1 Mathematical Concepts for Optimality and Compromise 245
11.1.2 Metabolism Is Shaped by Optimality 248
11.1.3 Optimality Approaches in Metabolic Modeling 250
11.1.4 Metabolic Strategies 252
11.1.5 Optimal Metabolic Adaptation 253
11.2 Optimal Enzyme Concentrations 255
11.2.1 Optimization of Catalytic Properties of Single Enzymes 255
11.2.2 Optimal Distribution of Enzyme Concentrations in a Metabolic Pathway 257
11.2.3 Temporal Transcription Programs 259
11.3 Evolution and Self-Organization 261
11.3.1 Introduction 261
11.3.2 Selection Equations for Biological Macromolecules 263
11.3.3 The Quasispecies Model: Self-Replication with Mutations 265
Further Reading 267

Contents vii
15 Mathematical and Physical Concepts 381
15.1 Linear Algebra 381
15.1.1 Linear Equations 381
15.1.2 Matrices 384
15.2 Dynamic Systems 386
15.2.1 Describing Dynamics with Ordinary Differential Equations 386
15.2.2 Linearization of Autonomous Systems 388
15.2.3 Solution of Linear ODE Systems 388
15.2.4 Stability of Steady States 388
15.2.5 Global Stability of Steady States 390
15.2.6 Limit Cycles 390
15.3 Statistics 391
15.3.1 Basic Concepts of Probability Theory 391
15.3.2 Descriptive Statistics 396
15.3.3 Testing Statistical Hypotheses 399
15.3.4 Linear Models 401
15.3.5 Principal Component Analysis 404
15.4 Stochastic Processes 405
15.4.1 Chance in Physical Theories 405
15.4.2 Mathematical Random Processes 406
15.4.3 Brownian Motion as a Random Process 407
15.4.4 Markov Processes 409
15.4.5 Markov Chains 410
15.4.6 Jump Processes in Continuous Time 410
15.4.7 Continuous Random Processes 411
15.4.8 Moment-Generating Functions 412
15.5 Control of Linear Dynamical Systems 412
15.5.1 Linear Dynamical Systems 413
15.5.2 System Response and Linear Filters 414
15.5.3 Random Fluctuations and Spectral Density 415
15.5.4 The Gramian Matrices 415
15.5.5 Model Reduction 416
15.5.6 Optimal Control 416
15.6 Biological Thermodynamics 417
15.6.1 Microstate and Statistical Ensemble 417
15.6.2 Boltzmann Distribution and Free Energy 418
15.6.3 Entropy 419
15.6.4 Equilibrium Constant and Energies 421
15.6.5 Chemical Reaction Systems 422
15.6.6 Nonequilibrium Reactions 424
15.6.7 The Role of Thermodynamics in Systems Biology 425
15.7 Multivariate Statistics 426
15.7.1 Planning and Designing Experiments for Case-Control Studies 426
15.7.2 Tests for Differential Expression 427
15.7.3 Multiple Testing 428
15.7.4 ROC Curve Analysis 429
15.7.5 Clustering Algorithms 430
15.7.6 Cluster Validation 435
15.7.7 Overrepresentation and Enrichment Analyses 436
15.7.8 Classification Methods 438
Exercises 441
References 443

16 Databases 445
16.1 General-Purpose Data Resources 445
16.1.1 PathGuide 445
16.1.2 BioNumbers 446
16.2 Nucleotide Sequence Databases 446
16.2.1 Data Repositories of the National Center for Biotechnology Information 446
16.2.2 GenBank/RefSeq/UniGene 446
16.2.3 Entrez 447
16.2.4 EMBL Nucleotide Sequence Database 447
16.2.5 European Nucleotide Archive 447
16.2.6 Ensembl 447
16.3 Protein Databases 448
16.3.1 UniProt/Swiss-Prot/TrEMBL 448
16.3.2 Protein Data Bank 448
16.3.3 PANTHER 448
16.3.4 InterPro 448
16.3.5 iHOP 449
16.4 Ontology Databases 449
16.4.1 Gene Ontology 449
16.5 Pathway Databases 449
16.5.1 KEGG 450
16.5.2 Reactome 450
16.5.3 ConsensusPathDB 451
16.5.4 WikiPathways 451
16.6 Enzyme Reaction Kinetics Databases 451
16.6.1 BRENDA 451
16.6.2 SABIO-RK 452
16.7 Model Collections 452
16.7.1 BioModels 452
16.7.2 JWS Online 452
16.8 Compound and Drug Databases 452
16.8.1 ChEBI 453
16.8.2 Guide to PHARMACOLOGY 453
16.9 Transcription Factor Databases 453
16.9.1 JASPAR 453
16.9.2 TRED 453
16.9.3 Transcription Factor Encyclopedia 454
16.10 Microarray and Sequencing Databases 454
16.10.1 Gene Expression Omnibus 454
16.10.2 ArrayExpress 454
References 455
<table>
<thead>
<tr>
<th>17</th>
<th>Software Tools for Modeling</th>
<th>457</th>
</tr>
</thead>
<tbody>
<tr>
<td>17.1</td>
<td>13C-Flux2</td>
<td>458</td>
</tr>
<tr>
<td>17.2</td>
<td>Antimony</td>
<td>458</td>
</tr>
<tr>
<td>17.3</td>
<td>Berkeley Madonna</td>
<td>459</td>
</tr>
<tr>
<td>17.4</td>
<td>BIOCHAM</td>
<td>459</td>
</tr>
<tr>
<td>17.5</td>
<td>BioNetGen</td>
<td>459</td>
</tr>
<tr>
<td>17.6</td>
<td>Biopython</td>
<td>459</td>
</tr>
<tr>
<td>17.7</td>
<td>BioTapestry</td>
<td>460</td>
</tr>
<tr>
<td>17.8</td>
<td>BioUML</td>
<td>460</td>
</tr>
<tr>
<td>17.9</td>
<td>CellDesigner</td>
<td>460</td>
</tr>
<tr>
<td>17.10</td>
<td>CellNetAnalyzer</td>
<td>460</td>
</tr>
<tr>
<td>17.11</td>
<td>Copasi</td>
<td>461</td>
</tr>
<tr>
<td>17.12</td>
<td>CPN Tools</td>
<td>461</td>
</tr>
<tr>
<td>17.13</td>
<td>Cytoscape</td>
<td>461</td>
</tr>
<tr>
<td>17.14</td>
<td>E-Cell</td>
<td>461</td>
</tr>
<tr>
<td>17.15</td>
<td>EvA2</td>
<td>461</td>
</tr>
<tr>
<td>17.16</td>
<td>FEniCS Project</td>
<td>462</td>
</tr>
<tr>
<td>17.17</td>
<td>Genetic Network Analyzer (GNA)</td>
<td>462</td>
</tr>
<tr>
<td>17.18</td>
<td>Jarnac</td>
<td>462</td>
</tr>
<tr>
<td>17.19</td>
<td>JDesigner</td>
<td>463</td>
</tr>
<tr>
<td>17.20</td>
<td>JSim</td>
<td>463</td>
</tr>
<tr>
<td>17.21</td>
<td>KNIME</td>
<td>463</td>
</tr>
<tr>
<td>17.22</td>
<td>libSBML</td>
<td>464</td>
</tr>
<tr>
<td>17.23</td>
<td>MASON</td>
<td>464</td>
</tr>
<tr>
<td>17.24</td>
<td>Mathematica</td>
<td>464</td>
</tr>
<tr>
<td>17.25</td>
<td>MathSBML</td>
<td>465</td>
</tr>
<tr>
<td>17.26</td>
<td>Matlab</td>
<td>465</td>
</tr>
<tr>
<td>17.27</td>
<td>MesoRD</td>
<td>465</td>
</tr>
<tr>
<td>17.28</td>
<td>Octave</td>
<td>465</td>
</tr>
<tr>
<td>17.29</td>
<td>Omix Visualization</td>
<td>466</td>
</tr>
<tr>
<td>17.30</td>
<td>OpenCOR</td>
<td>466</td>
</tr>
<tr>
<td>17.31</td>
<td>Oscill8</td>
<td>466</td>
</tr>
<tr>
<td>17.32</td>
<td>PhysioDesigner</td>
<td>466</td>
</tr>
<tr>
<td>17.33</td>
<td>PottersWheel</td>
<td>467</td>
</tr>
<tr>
<td>17.34</td>
<td>PyBioS</td>
<td>467</td>
</tr>
<tr>
<td>17.35</td>
<td>PySCeS</td>
<td>467</td>
</tr>
<tr>
<td>17.36</td>
<td>R</td>
<td>468</td>
</tr>
<tr>
<td>17.37</td>
<td>SAAM II</td>
<td>468</td>
</tr>
<tr>
<td>17.38</td>
<td>SBMLeditor</td>
<td>468</td>
</tr>
<tr>
<td>17.39</td>
<td>SemanticSBML</td>
<td>468</td>
</tr>
<tr>
<td>17.40</td>
<td>SBML-PET-MPI</td>
<td>469</td>
</tr>
<tr>
<td>17.41</td>
<td>SBMLsimulator</td>
<td>469</td>
</tr>
<tr>
<td>17.42</td>
<td>SBMLSqueezer</td>
<td>469</td>
</tr>
<tr>
<td>17.43</td>
<td>SBML Toolbox</td>
<td>470</td>
</tr>
<tr>
<td>17.44</td>
<td>SBtoolbox2</td>
<td>470</td>
</tr>
<tr>
<td>17.45</td>
<td>SBML Validator</td>
<td>470</td>
</tr>
<tr>
<td>17.46</td>
<td>SensA</td>
<td>470</td>
</tr>
<tr>
<td>17.47</td>
<td>SmartCell</td>
<td>471</td>
</tr>
<tr>
<td>17.48</td>
<td>STELLA</td>
<td>471</td>
</tr>
<tr>
<td>17.49</td>
<td>STEPS</td>
<td>471</td>
</tr>
<tr>
<td>17.50</td>
<td>StochKit2</td>
<td>471</td>
</tr>
<tr>
<td>17.51</td>
<td>SystemModeler</td>
<td>472</td>
</tr>
<tr>
<td>17.52</td>
<td>Systems Biology Workbench</td>
<td>472</td>
</tr>
<tr>
<td>17.53</td>
<td>Taverna</td>
<td>472</td>
</tr>
<tr>
<td>17.54</td>
<td>VANTED</td>
<td>473</td>
</tr>
<tr>
<td>17.55</td>
<td>Virtual Cell (VCell)</td>
<td>473</td>
</tr>
<tr>
<td>17.56</td>
<td>xCellerator</td>
<td>473</td>
</tr>
<tr>
<td>17.57</td>
<td>XPPAUT</td>
<td>473</td>
</tr>
</tbody>
</table>

Exercises 474

References 474

Index 475