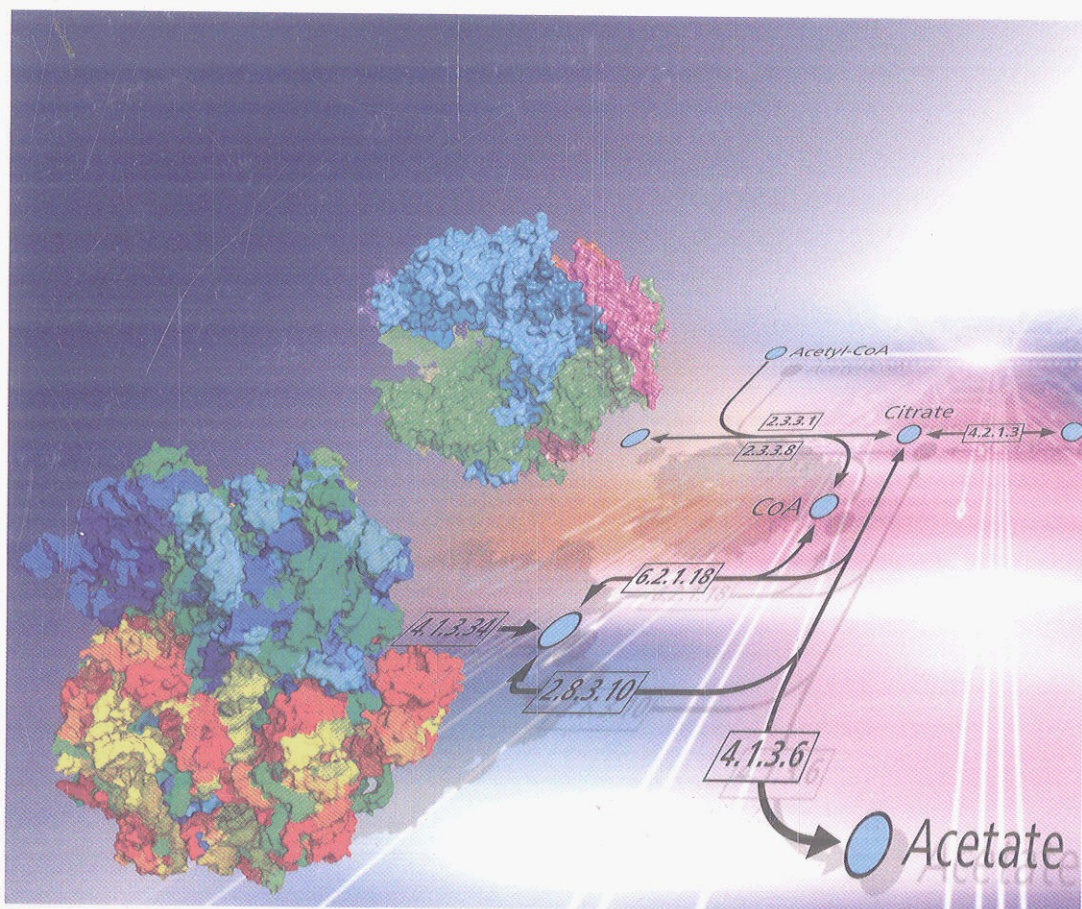


Volkhard Helms

WILEY-VCH

# Principles of Computational Cell Biology

From Protein Complexes to Cellular Networks



# Contents

## Preface XI

<b>1</b>	<b>Networks in Biological Cells</b>	<b>1</b>
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 Network Model	3
1.2	Biological Background	4
1.2.1	Cellular Components	6
1.2.2	Spatial Organization of Eukaryotic Cells – Compartments	7
1.2.3	Cellular Organisms	7
1.3	Cellular Pathways	7
1.3.1	Biochemical Pathways	7
1.3.2	Enzymatic Reactions	8
1.3.3	Signal Transduction	11
1.3.4	Cell Cycle	11
1.4	Ontologies and Databases	12
1.4.1	Ontologies	12
1.4.2	Systems Biology Markup Language	12
1.4.3	KEGG	13
1.4.4	Brenda	13
1.5	Methods in Cellular Modeling	14
<b>2</b>	<b>Algorithms on Mathematical Graphs</b>	<b>17</b>
2.1	Primer on Mathematical Graphs	17
2.2	A Few Words about Algorithms and Computer Programs	18
2.2.1	Implementation of Algorithms	19
2.2.2	Classes of Algorithms	20
2.3	Data Structures for Graphs	21

2.4	Dijkstra's Algorithm	23
2.4.1	Description of the Algorithm	25
2.4.2	Pseudocode	27
2.4.3	Running Time	29
2.5	Minimum Spanning Tree	29
2.5.1	Kruskal's Algorithm	31
2.6	Graph Drawing	31
<b>3</b>	<b>Protein-Protein Interaction Networks – Pairwise Connectivity</b>	<b>39</b>
3.1	Principles of Protein-Protein Interactions	39
3.2	Experimental High-Throughput Methods for Detecting Protein-Protein Interactions	40
3.2.1	Gel Electrophoresis	41
3.2.2	Two-Dimensional Gel Electrophoresis	41
3.2.3	Affinity Chromatography	42
3.2.4	Yeast Two-Hybrid Screening	42
3.2.5	Synthetic Lethality	44
3.2.6	Gene Coexpression	44
3.2.7	Mass Spectroscopy	44
3.2.8	Databases for Interaction Networks	44
3.2.9	Overlap of Interactions	45
3.2.10	Criteria to Judge the Reliability of Interaction Data	47
3.2.11	How Many Protein-Protein Interactions can be Expected in Yeast?	48
3.3	Bioinformatic Prediction of Protein-Protein Interactions	49
3.3.1	Analysis of Gene Order	49
3.3.2	Phylogenetic Profiling/Coevolutionary Profiling	50
3.3.3	Coevolution	51
3.4	Bayesian Networks for Judging the Accuracy of Interactions	52
3.4.1	Bayes' Theorem	53
3.4.2	Bayesian Network	54
3.4.3	Application of Bayesian Networks to Protein-Protein Interaction Data	55
3.4.3.1	Measurement of reliability "likelihood ratio"	55
3.4.3.2	Prior and posterior odds	56
3.4.3.3	A worked example: parameters of the naïve Bayesian network for essentiality	57
3.4.3.4	Fully connected experimental network	57
3.5	Protein Domain Networks	59
<b>4</b>	<b>Protein-Protein Interaction Networks – Structural Hierarchies</b>	<b>67</b>
4.1	Protein Interaction Graph Networks	67
4.1.1	Degree Distribution	68
4.1.2	Clustering Coefficient	69
4.2	Finding Cliques	71

4.3	Random Graphs	72
4.4	Scale-Free Graphs	73
4.5	Detecting Communities in Networks	75
4.5.1	Divisive Algorithms for Mapping onto Tree	78
4.6	Modular Decomposition	82
4.6.1	Modular Decomposition of Graphs	82
4.7	Network Growth Mechanisms	86
<b>5</b>	<b>Gene Regulatory Networks</b>	<b>99</b>
5.1	Regulation of Gene Transcription at Promoters	100
5.2	Gene Regulatory Networks	101
5.2.1	Gene Regulatory Network of <i>E. coli</i>	101
5.3	Graph Theoretical Models	105
5.3.1	Cocexpression Networks	105
5.3.2	Bayesian Networks	106
5.4	Dynamic Models	106
5.4.1	Boolean Networks	106
5.4.2	Reverse Engineering Boolean Networks	107
5.4.3	Differential Equations Models	110
5.5	Motifs	111
5.5.1	Feed-Forward Loop (FFL)	112
5.5.2	SIM Motif	112
5.5.3	Densely Overlapping Region (DOR)	112
<b>6</b>	<b>Metabolic Networks</b>	<b>115</b>
6.1	Introduction	115
6.2	Stoichiometric Matrix	118
6.3	Linear Algebra Primer	121
6.3.1	Matrices: Definitions and Notations	121
6.3.2	Adding, Subtracting and Multiplying Matrices	121
6.3.3	Linear Transformations, Ranks and Transpose	122
6.3.4	Square Matrices and Matrix Inversion	123
6.3.5	Eigenvalues of Matrices	124
6.3.6	System of Linear Equations	124
6.4	Flux Balance Analysis	125
6.5	Double Description Method	128
6.6	Extreme Pathways and Elementary Modes	133
6.6.1	Analysis of Extreme Pathways	137
6.6.2	Elementary Flux Modes	139
6.7	Minimal Cut Sets	140
6.7.1	Applications of Minimal Cut Sets	144
6.8	High-Flux Backbone	146
<b>7</b>	<b>Kinetic Modeling of Cellular Processes</b>	<b>155</b>
7.1	Ordinary Differential Equation Models	155

7.1.1	Examples for ODEs	156
7.2	Modeling Cellular Feedback Loops by ODEs	158
7.2.1	Protein Synthesis and Degradation: Linear Response	159
7.2.2	Phosphorylation/Dephosphorylation – Hyperbolic Response	160
7.2.3	Phosphorylation/Dephosphorylation – Buzzer	162
7.2.4	Perfect Adaptation – Sniffer	163
7.2.5	Positive Feedback – One-Way Switch	164
7.2.6	Mutual Inhibition – Toggle Switch	165
7.2.7	Negative Feedback – Homeostasis	166
7.2.8	Negative Feedback: Oscillatory Response	166
7.2.9	Cell Cycle Control System	167
7.3	Partial Differential Equations	169
7.3.1	Spatial Gradients of Signaling Activities	170
7.4	Dynamic Monte Carlo (Gillespie Algorithm)	172
7.4.1	Basic Outline of the Gillespie Method	173
7.5	Stochastic Modeling of a Small Molecular Network	173
7.5.1	Model System: Bacterial Photosynthesis	174
7.5.2	Pools-and-Proteins Model	176
7.5.3	Evaluating the Binding and Unbinding Kinetics	177
7.5.4	Pools of the Chromatophore Vesicle	178
7.5.5	Results for the Steady-State Regimes of the Vesicle	179
7.6	Parameter Optimization with Genetic Algorithms	182

## **8 Structures of Protein Complexes and Subcellular Structures** 193

8.1	Examples of Protein Complexes	193
8.2	Complexome of <i>S. cerevisiae</i>	197
8.3	Experimental Determination of Three-dimensional Structures of Protein Complexes	199
8.3.1	X-ray Crystallography	199
8.3.2	NMR	200
8.3.3	Electron Crystallography/Electron Microscopy	201
8.3.4	Immuno-electron Microscopy	201
8.3.5	Fluorescence Resonance Energy Transfer	202
8.4	Density Fitting	204
8.4.1	Correlation-based Fitting	204
8.5	Fourier Transformation	206
8.5.1	Fourier Series	206
8.5.2	Continuous Fourier Transform	207
8.5.3	Discrete Fourier Transform	207
8.5.4	Convolution Theorem	208
8.5.5	Fast Fourier Transformation	208
8.6	Advanced Density Fitting	210
8.6.1	Laplacian Filter	211
8.6.2	Fitting Using Core Downweighting	212
8.6.3	Core-weighted Correlation Function	214

8.6.4	Surface Overlap Maximization (SOM)	215
8.7	FFT Protein-Protein Docking	216
8.8	Prediction of Assemblies from Pairwise Docking	218
8.9	Electron Tomography	221
8.9.1	Reconstruction of a Phantom Cell	222
<b>9</b>	<b>Biomolecular Association and Binding</b>	<b>231</b>
9.1	Modeling by Homology	231
9.2	Structural Properties of Protein-Protein Interfaces	233
9.2.1	Size and Shape	233
9.2.2	Hot Spots	235
9.2.3	An Experimental Model System: Human Growth Hormone and its Receptor	236
9.3	Bioinformatic Prediction of Protein-Protein Interfaces	239
9.3.1	Amino acid Composition of Protein Interfaces	239
9.3.2	Pairing Propensities	240
9.3.3	Interface Statistical Potentials	240
9.3.4	Conservation at Protein Interfaces	241
9.3.5	Correlated Mutations at Protein Interfaces	243
9.3.6	Classification of Protein Interfaces	245
9.4	Forces Important for Biomolecular Association	246
9.5	Protein-Protein Association	249
9.5.1	Brownian Dynamics Simulations	250
9.6	Assembly of Macromolecular Complexes: the Ribosome	254
<b>10</b>	<b>Integrated Networks</b>	<b>261</b>
10.1	Correlating Interactome and Gene Regulation	261
10.2	Response of Gene Regulatory Network to Outside Stimuli	263
10.3	Integrated Analysis of Metabolic and Regulatory Networks	266
<b>11</b>	<b>Outlook</b>	<b>271</b>
	<b>Index</b>	<b>273</b>