Cheminformatics and Bioinformatics at the Interface with Systems Biology

Bridging Chemistry and Medicine

By

Aman Chandra Kaushik
Shanghai Jiao Tong University, China
Email:amanbioinfo@gmail.com

Aamir Mehmood
Shanghai Jiao Tong University, China
Email:aamirmehmood@sjtu.edu.cn

Dong-Qing Wei
Shanghai Jiao Tong University, China
Email:dqwei@sjtu.edu.cn

Sadia Nawab
Shanghai Jiao Tong University, China
Email:sadianawab@sjtu.edu.cn

Shakti Sahi
Gautam Buddha University, India
Email:shaktis@gbu.ac.in

and

Ajay Kumar
National Sun Yat-sen University, Taiwan
Email:ajaynsysu@mem.nsysu.edu.tw

ROYAL SOCIETY OF CHEMISTRY

Contents

Chapter 1	**Systems Biology and Drug Target Identification**	**1**
	1.1 Introduction	1
	1.2 Systems Biology and Drug Targets	2
	1.2.1 Systems-based Drug Identification	3
	1.2.2 Disease as a Network State	3
	1.2.3 Regulating Network Dynamics	3
	1.2.4 Allosteric Sites as Drug Binding Targets	4
	1.2.5 Controlling Drug Target Gene Expression	5
	1.2.6 Impact of Drug Binding Kinetics	6
	1.3 Novel Guidelines for Structure-based Drug Design (SBDD) in the Systems Biology Era	7
	1.3.1 *De Novo* Drug Design	7
	1.3.2 Multi-target Drug Design	8
	1.3.3 Repositioning Drug Analogs	9
	1.3.4 Drug Design for Targeting Protein–Protein Interactions (PPIs)	10
	References	10
Chapter 2	**Modulating Drug Target Gene Expression in Cancer**	**16**
	2.1 Introduction	16
	2.2 Database Sources of Cancer-related Genes	18
	2.2.1 Based on the Knowledge of Pathways	19
	2.3 Network-based Methods	20

Theoretical and Computational Chemistry Series No. 24
Cheminformatics and Bioinformatics at the Interface with Systems Biology: Bridging Chemistry and Medicine
By Aman Chandra Kaushik, Aamir Mehmood, Dong-Qing Wei, Sadia Nawab, Shakti Sahi and Ajay Kumar

© Aman Chandra Kaushik, Aamir Mehmood, Dong-Qing Wei, Sadia Nawab, Shakti Sahi and Ajay Kumar 2024

Published by the Royal Society of Chemistry, www.rsc.org

2.4 Combinatorial Patterns of Cancer Mutations 22
2.4.1 Data Integration and Methodological Combination 23
2.4.2 Subtypes of Cancer and Pan-cancer Analysis 24
2.5 Bias Towards Prior Knowledge 24
2.6 Drawbacks of Protein Network Interactions 25
2.7 Validation of Experiments 25
2.8 Conclusion 26
References 26

Chapter 3 Drug Repositioning Using Genome-wide Screening and Systems Biology Approaches and Applications 31
3.1 Introduction 31
3.2 High-performance Identification of Drug Targets and Repositionings 33
3.3 DBs for Bioinformatics and *In Vitro*-based Assays 33
3.3.1 Computational-based Assay DBs 34
3.3.2 Genome-based Search for Drug Rescuing 36
3.3.3 Drug Target Selection Based on RNAi Search 36
3.3.4 DTI Search *Via* Genome-based CRISPR-Cas9 Analysis 36
3.4 Applications of Systems Biology in Drug Repositioning 37
3.4.1 Network-dependent Approaches for Drug Relocation 37
3.5 Role of Machine Learning in Drug Relocations 39
3.6 Cheminformatics in Drug Relocation 40
3.7 Computational Approaches for Drug Repositioning 41
3.8 Repositioning Drugs Through Signature-based Approaches 41
3.9 Conclusion 42
References 43

Chapter 4 New Directions in Systems Biology-based Target Identification and Cancer Genome Analysis 52
4.1 Introduction 52
4.2 Prediction of Drug–Target Communications Using Systems Biology 53
4.3 Developments in Targeted Screening and Drug Design Using Systems Biology 54
4.4 Drug Design *Via* High-throughput Approaches 55
4.5 Target Identification Through Molecular Modeling and Systems Biology 56
4.6 Developing Drug–Target Networks Using Systems Biology 58
4.7 Predicting Drug–Target Interactions Through Systems Biology 59

4.8 Target Confirmation *Via* Systems Biology Methods 60
4.9 Ways of Studying and Plotting Cancer Genome Information 62
4.10 Understanding the Cancer Genome Through *In Silico* Approaches 62
4.10.1 Discovery of Driver Genes in Cancer 62
4.10.2 A Set of Mutational Genes in Cancer 63
4.11 Understanding Cancer's Genetic Foundation 64
4.12 Mutational Patterns Revealing the Key Driver 65
4.13 Estimating the Impact of a Missense Mutation in Cancer 67
4.14 Understanding Cancer Gene Networks 68
4.15 Concluding Remarks and Future Perspectives 69
References 70

Chapter 5 Molecular Investigation of Protein–Protein Interaction Candidates Related to the Mammalian Brain 81
5.1 Introduction 81
5.2 Parkinson's Disease 82
5.3 Alzheimer's Disease – Another Form of Dementia 83
5.4 PPI Data Sets for Novel Biological Insights 85
5.5 Understanding Parkin (PARK2) 86
5.6 Systems Biology 86
5.6.1 Biochemical Pathway Analysis 86
5.6.2 Alzheimer's Disease from Systems Biology and Gene Networks Perspectives 87
5.6.3 Administrative Catalogs 88
5.6.4 Mutations and "Omics" Channels in AD 88
5.6.5 AD's Transcriptome Channels 89
5.6.6 Protein Communications in AD 91
5.6.7 Protein Communication Channels and Multifaceted Neurodegenerative Diseases 92
5.6.8 Parkinson's Disease Scenario 92
5.6.9 Using Protein Communications to Understand Gene Interactomes 94
5.6.10 Comparing PD with Other Neurodegenerative Conditions 95
5.7 Limitations and Future Directions 96
5.8 Future Perspectives and Conclusion 98
References 98

Chapter 6 Biological Systems to Computational Systems Biology 108
6.1 Introduction 108
6.2 Quantitative Computational Models of Molecular Self-assembly in Systems Biology 112
6.3 Validating the Importance of Self-assembly 114

6.3.1 The Role of Self-assembly in General Cell Biology 114
6.4 Modeling and Simulation of a Self-assembly 116
6.4.1 The Challenge of Quantitative Modeling of Self-assembly Reaction Networks 116
6.5 Modeling Approaches 117
6.5.1 Models for the Mass Action Differential Equation (DE) 118
6.5.2 Structures Concerning Brownian Dynamics (BD) 118
6.5.3 Approaches for Hypothetical Simulation Systems 119
6.6 Self-assembly in a Wider Context 121
6.7 Review of Machine Learning Strategies for Studying Multi-omics 122
6.7.1 Background 122
6.7.2 Hurdles While Studying Machine Learning-based Multi-omics 124
6.8 Final Thoughts 125
References 126

Chapter 7 Controlled Vocabularies and Semantics in Systems Biology 136
7.1 Introduction 136
7.2 Typical Structure of SBO 139
7.3 Simulation Workflows: KiSAO 141
7.4 Arithmetic Outcomes: TEDDY 143
7.5 Modeling and Simulation of Workflow Ontologies 144
7.6 Conclusion 145
References 146

Chapter 8 Single-molecule Imaging in Biosystems 150
8.1 Introduction 150
8.2 Various Methods of Cellular SMI 151
8.2.1 Microscopy 151
8.2.2 Procedural Shortcomings of SMI 152
8.3 Applications of SMI in Systems Biology 153
8.3.1 Calculating the Molecular Ratio in Cellular Responses 153
8.3.2 Dynamic Breakdown of Living Cells 154
8.3.3 Coupling Numerical Modeling with SMIs 155
8.4 Kinetics of a Solo Molecule 155
8.5 SMI in Biosystems (Within the Nucleus) 156
8.5.1 DNA 157
8.5.2 mRNA 158
8.5.3 Proteins 159

8.6 Major Drawbacks 160
8.7 Assumptions and Viewpoints 161
8.8 Imaging the Ubiquitination of Mammalian Cells 162
8.8.1 Methods to Study Ubiquitination 162
8.9 Conclusion 168
References 168

Chapter 9 Tracking the Emergence of Synthetic Biology 176
9.1 The Early Years of Synthetic Biology 176
9.2 Creating Standard Genetic Parts/Circuits to Fabricate Natural Organisms with Novel Functions 176
9.3 Synthetically Designed Genomes 177
9.4 Recent Developments in Synthetic Biology 178
9.4.1 Apparatus 178
9.4.2 Production of Biopharmaceuticals 179
9.4.3 Immunotherapy 179
9.4.4 A Maintainable Biochemical Commerce 179
9.4.5 Biosensing and Bioremediation 180
9.4.6 Financing Synthetic Biology 180
9.5 Potential Dangers in Synthetic Biology 181
9.5.1 Biosafety Alarms 181
9.5.2 Biosecurity Alarms 183
9.5.3 Moral Apprehensions 184
9.5.4 Precautionary Procedures at the Laboratory Level for Synthetic Biology-related Safety and Security 185
9.6 Synthetic Biology-related Regulatory Policy 188
9.6.1 Scientists' Code of Conduct 188
9.6.2 Governance at the National Level 189
9.6.3 International Societies' Efforts 190
9.7 Conclusions 190
References 191

Chapter 10 Synthetic Biology: Fostering the Cyber-biological Revolution 196
10.1 Introduction 196
10.2 DNA as the New Silicon 197
10.3 Model-powered Life Cycle Development 198
10.4 *"The Times They Are a-Changin'"* 200
10.5 The Industrial Revolution in Cyber-biology 200
10.6 Computational Pathology (CPATH) 202
10.6.1 History and Promises of Computational Pathology 203
10.6.2 Correlating Images with Patient Outcome 204
10.6.3 Helping with the Diagnosis 204
10.6.4 Finding Unusual Characteristics 205

10.6.5 Hurdles and Solutions for Implementing Computational Pathology 205
10.6.6 Data Variability 207
10.6.7 Public Sources 208
10.6.8 Crowdsourcing 208
10.6.9 Active Learning 208
10.6.10 Quality Control and Reliability of the Algorithm 209
10.6.11 Understanding Algorithms 209
10.7 Ethics 210
10.8 Cyber-security 210
10.9 Conclusion and Future Recommendations 211
10.9.1 Sensing the Future of Bioinformational Engineering 212
References 214

Chapter 11 Computational Systems Chemical Biology 222
11.1 Overview 222
11.2 SCB Existing Repositories, Assembly, and Operations 224
11.2.1 Complex Bioactivity Databases 226
11.2.2 Pathway-specific Databases 228
11.2.3 Bioavailability Databases 229
11.2.4 Drugs, Targets, and Clinical Implications Databases 230
11.3 Modeling SCB Information to Predict Drug–Target Associations: Computational Approaches 231
11.4 Biological Network Simulations 234
11.4.1 Pathway Simulation 234
11.5 Integrated Cheminformatics and Network Simulations 235
11.6 Conclusion 236
References 236

Chapter 12 Systems Biology Methods to Assist in the Discovery of Efficacious Drug Combinations 239
12.1 Introduction 239
12.2 Drug Combinations with Synergistic Effects Being Quantified 240
12.2.1 Loewe Additivity 242
12.2.2 Bliss Independence 242
12.3 Signaling Network Computational Models 243
12.3.1 Mass Action- and Enzyme Kinetics-based Models 243
12.3.2 Logic-based Models 244
12.3.3 Approach to Normalized Hill Differential Equation Modeling 245

12.3.4 Statistical Association-based Modeling Approach 246
12.4 Approaches for Predicting Drug Combinations Based on Signature 246
12.4.1 Prediction of Drug Combinations Using Network-based Approaches 247
12.4.2 Integrating Functional Genomics and Computational Methods for Identifying Drug Combinations 249
12.4.3 High-throughput Drug Combination Screens 250
12.5 Conclusion 252
References 252

Subject Index **258**