Contents

List of Contributors  XV
Foreword     XXI
Preface      XXIII

1  De Novo Design: From Models to Molecules  1
   Gisbert Schneider and Karl-Heinz Baringhaus
   1.1 Molecular Representation  1
   1.2 The Molecular Design Cycle  9
   1.3 Receptor–Ligand Interaction  14
   1.4 Modeling Fitness Landscapes  21
   1.4.1 Naïve Bayes Classifier  26
   1.4.2 Artificial Neural Network  27
   1.4.3 Support Vector Machine  27
   1.4.4 Gaussian Process  28
   1.5 Strategies for Compound Construction  30
   1.6 Strategies for Compound Scoring  33
   1.6.1 Receptor-Based Scoring  35
   1.6.2 Ligand-Based Scoring  37
   1.7 Flashback Forward: A Brief History of De Novo Drug Design  37
   1.8 Conclusions  43
   Acknowledgments  43
   References  44

2  Coping with Complexity in Molecular Design  57
   Michael M. Hann and Andrew R. Leach
   2.1 Introduction  57
   2.2 A Simple Model of Molecular Interactions  58
   2.3 Enhancements to the Simple Complexity Model  60
   2.4 Enumerating and Sampling the Complexity of Chemical Space  61
   2.5 Validation of the Complexity Model  65
   2.6 Reductionism and Drug Design  67
Contents

2.7 Complexity and Information Content as a Factor in De Novo Design 69
2.8 Complexity of Thermodynamic Entropy and Drug Design 73
2.9 Complex Systems, Emergent Behavior, and Molecular Design 74
Acknowledgments 75
References 75

3 The Human Pocketome 79
Ruben Abagyan and Clarisse Gravina Ricci
3.1 Predicted Pockets 79
3.2 Compilation of the Validated Human Pocketome 83
3.3 Diversity and Redundancy of the Human Pocketome 85
3.4 Compound Activity Prediction by Ligand-Pocket Docking and Scoring 87
3.4.1 Optimizing Pocket Sets for Reliable Docking and Scoring Results 87
3.4.2 Difficult Cases: Unusually Large and Multifunctional Pockets 88
3.5 Pocketome-Derived 3D Chemical Fields as Activity Prediction Models 90
3.6 Clustering the Ligands by Function and Subpockets 92
3.7 Conclusions 94
Acknowledgments 94
References 94

4 Structure-Based De Novo Drug Design 97
Alla Srinivas Reddy, Lu Chen, and Shuxing Zhang
4.1 Introduction 97
4.2 Current Progress in SBDND Methodologies 99
4.2.1 Identification of Binding Site 100
4.2.2 Design of Molecules 101
4.2.2.1 Atom-Based versus Fragment-Based Methods 101
4.2.2.2 Pharmacophore-Based Methods 105
4.2.3 Searching the Chemical Space 105
4.2.3.1 Monte Carlo-Based Methods 106
4.2.3.2 Evolutionary Algorithms 106
4.2.4 Scoring Methods 108
4.2.4.1 Force-Field-Based Scoring Functions 108
4.2.4.2 Empirical Scoring Functions 109
4.2.4.3 Knowledge-Based Scoring Functions 109
4.2.4.4 Consensus Scoring 109
4.2.5 Synthetic Accessibility 110
4.3 Recent Applications of Structure-Based De Novo Design 110
4.4 Perspectives and Conclusion 115
Acknowledgment 116
References 116
5 De Novo Design by Fragment Growing and Docking  

Jacob D. Durrant and Rommie E. Amaro

5.1 Introduction 125
5.2 Case Study I: High-Throughput Screening with Dr Feils 126
5.2.1 Target Identification 126
5.2.2 Small-Molecule Library Design 126
5.2.2.1 Computer Docking 128
5.2.2.2 Pharmacophore Searching 129
5.2.3 High-Throughput Screening 129
5.2.4 Optimization 130
5.3 Case Study II: Fragment-Based Drug Design with Dr Goode 130
5.3.1 Library Generation 130
5.3.1.1 Computational Techniques for Library Refinement 132
5.3.2 Detection Methods 132
5.3.2.1 Functional/High-Concentration Screening 132
5.3.2.2 Fluorescence-Based Thermal Shift Assay (TSA) 133
5.3.2.3 Surface Plasmon Resonance (SPR) 133
5.3.2.4 Mass Spectrometry (MS) 133
5.3.2.5 Nuclear Magnetic Resonance (NMR) 133
5.3.2.6 X-Ray Crystallography 134
5.3.3 Screening 134
5.3.4 Optimization 135
5.3.5 Final Products 137
5.4 Conclusion 138

Disclaimer 138
Acknowledgments 138
References 138

6 Hit and Lead Identification from Fragments 143

Michael Mazanetz, Richard Law, and Mark Whittaker

6.1 Introduction to FBDD 144
6.2 Fragment Library Design Incorporating Computational Methods 148
6.2.1 Fragment Library Design Strategies 148
6.2.2 Molecular Attributes and Physicochemical Properties 150
6.2.3 Influence of Screening Method on Library Selection 151
6.2.4 Removal of Undesirable Functionality 152
6.2.5 Size of Library and Diversity 152
6.2.6 Focused Sets 154
6.2.7 Designing in Fragment Optimization 154
6.3 Fragment Screening 155
6.3.1 Screening by X-Ray Crystallography 155
6.3.2 Screening by NMR 156
6.3.3 Screening by SPR 157
6.3.4 Screening by Biochemical Assay 157
6.3.5 Thermal Shift Assays 158
6.3.6 Isothermal Titration Calorimetry (ITC) 160
6.3.7 Other Biophysical Assay Techniques 160
6.3.8 Assay Techniques for Membrane Proteins 161
6.3.9 Fragment Library Screening Using Computational Methods 161
6.3.10 Ligandability Screening Using Fragments 162
6.4 Fragment Prioritization for Optimization 165
6.4.1 Efficiency Metrics 165
6.4.2 Computational and Thermodynamic Methods for Fragment Selection and Prioritization 167
6.5 Fragment Hit Expansion and Fragment Evolution 170
6.6 Fragment Merging Principles 175
6.7 Fragment Linking Principles 177
6.8 Fragment-Assisted Drug Discovery (FADD) 182
6.9 Conclusion 183
Acknowledgments 184
References 184

7 Pharmacophore-Based De Novo Design 201
Wen-Jing Wang, Qi Huang, and Sheng-Yong Yang
7.1 Introduction 201
7.2 A Summary of the Algorithms of PhDD v1.0 202
7.2.1 The Basic Scheme of PhDD 202
7.2.2 Fragment and Linker Databases 203
7.2.3 Mapping of Fragments onto the Locations of Pharmacophore Features of the Pharmacophore Hypothesis 204
7.2.4 Connecting Fragments by Linkers 205
7.2.5 Assessments to the Generated Molecules 206
7.2.5.1 The Drug-Likeness Assessment 207
7.2.5.2 The Estimation of Bioactivity 207
7.2.5.3 The Assessment of Synthetic Accessibility 207
7.3 An Introduction to the Modifications in the Updated Version of PhDD (v2.0) 208
7.3.1 The Use of a Designated Fragment 209
7.3.2 Conformation Optimization in the Process of Molecular Construction 209
7.3.3 Two Pharmacophore Features Share One Fragment 210
7.4 Validation of PhDD 210
7.5 Concluding Remarks 212
Acknowledgment 213
References 213

8 3D-QSAR Approaches to De Novo Drug Design 215
Richard D. Cramer
8.1 Introduction 215
8.2 Current Methods 216
8.3 Leapfrog 217
8.4 Recent Advances 219
8.5 Conclusions 223
Acknowledgments 223
References 223

9 Ligand-Based Molecular Design Using Pseudoreceptors 227
Darren Fayne
9.1 Introduction 227
9.2 Pseudoreceptor Algorithms 231
9.3 Successful Applications Overview 232
9.4 Conclusions 240
Acknowledgments 241
References 241

10 Reaction-Driven De Novo Design: a Keystone for Automated Design of Target Family-Oriented Libraries 245
Markus Hartenfeller, Steffen Renner, and Edgar Jacoby
10.1 Introduction 245
10.2 Reaction-Driven Design: Tackling the Problem of Synthetic Feasibility 247
10.2.1 Exploiting the Valuable Knowledge Stored in Electronic Laboratory Notebooks 249
10.2.2 Assessing the Chemical Space of a Focused Set of Reactions 251
10.3 Successful Applications of Reaction-Driven De Novo Design 254
10.4 Reaction-Driven Design of Chemical Libraries Addressing Target Families 256
10.5 Conclusions 261
References 265

11 Multiobjective De Novo Design of Synthetically Accessible Compounds 267
Valerie J. Gillet, Michael J. Bodkin, and Dimitar Hristozov
11.1 Introduction 267
11.2 Design of Synthetically Accessible Compounds 269
11.3 Synthetic Accessibility Using Reaction Vectors 270
11.4 De Novo Design Using Evolutionary Algorithms 276
11.4.1 Optimizing Multiple Objectives 277
11.4.2 Multiobjective De Novo Design 279
11.4.3 Multiobjective De Novo Design Using Reaction Vectors 280
11.5 Conclusions 282
Acknowledgments 283
References 283
12 De Novo Design of Ligands against Multitarget Profiles 287
Jérémie Besnard and Andrew L. Hopkins
12.1 Introduction 287
12.2 Automating the Creativity of Ligand Design 289
12.3 Evolutionary Algorithm 294
12.4 Experimental Validation 295
12.5 Reducing Antitarget Activity 296
12.6 Optimizing D4 Receptor Potency 301
12.7 Designing Novel Ligands to a Defined Profile 301
12.8 Conclusion 304
Acknowledgments 306
References 306

13 Construction of Drug-Like Compounds by Markov Chains 311
Peter S. Kutchukian, Salla I. Virtanen, Eugen Lounkine, Meir Glick, and Eugene I. Shakhnovich
13.1 Introduction 311
13.2 FOG Algorithm and Library Generation 313
13.3 Applications 314
13.3.1 Overview 314
13.3.2 Target Class Prediction of FOG Compounds 314
13.3.3 Design of BACE-1 Inhibitors with FOG 316
13.4 Conclusion 319
Acknowledgments 320
References 320

14 Coping with Combinatorial Space in Molecular Design 325
Florian Lauck and Matthias Rarey
14.1 Introduction 325
14.2 Chemical Space 326
14.2.1 Size Estimation of Chemical Space 327
14.2.2 Enumeration of Chemical Subspaces 328
14.3 Combinatorial Space 330
14.3.1 Generation of Combinatorial Spaces 332
14.3.1.1 Combinatorial Space from Fragmentation 332
14.3.1.2 Computational Space from Chemical Knowledge 334
14.3.2 Manipulation of Combinatorial Space 335
14.3.3 Querying Combinatorial Spaces 336
14.3.3.1 Fragment Spaces 337
14.3.3.2 Reaction-Based Combinatorial Spaces 339
14.3.4 Other Applications of Combinatorial Space 340
14.3.5 Markush Structures 341
14.4 Visualization 342
14.5 Conclusion 343
References 343
15 Fragment-Based Design of Focused Compound Libraries 349
Uta Lessel
15.1 Introduction 349
15.2 General Workflow 351
15.3 Fragment Space 352
15.4 Query 355
15.5 FTrees Fragment Space Search 356
15.6 Scaffold Selection 356
15.7 Design of Focused Libraries 359
15.8 Application Example 360
15.9 Summary and Conclusions 366
Acknowledgments 367
References 367

16 Free Energy Methods in Ligand Design 373
Yvonne Westermaier and Roderick E. Hubbard
16.1 Free Energy (FE) Methods in Lead Optimization (LO) 373
16.1.1 FE Methods: An Emerging Tool in Industry? 374
16.1.2 Finding the Needle in a Haystack: The Role of FE Methods in Fine-Tuning Ligand Discovery 375
16.2 The Variety of In Silico Binding Affinity Methods 377
16.2.1 Thermodynamic Integration (TI) and Alchemical Transformations 377
16.2.2 Free Energy Perturbation (FEP) 378
16.2.3 Potential of Mean Force (PMF) Calculations 379
16.2.4 Nonequilibrium Approaches 380
16.2.5 Other MM-Based Methods 381
16.2.5.1 Linear Interaction Energy (LIE) 381
16.2.5.2 MM-PBSA and MM-GBSA 382
16.3 The Choice of a Method for Calculating Binding FE 382
16.3.1 MM-PBSA and MM-GBSA versus FEP/TI 383
16.3.2 LIE versus FEP/TI 383
16.3.3 PMF versus FEP 383
16.3.4 PMF versus TI 383
16.3.5 TI versus FEP 384
16.3.6 PMF/TI/FEP: Absolute or Relative Binding FEs? 384
16.3.7 Equilibrium versus Nonequilibrium Methods 385
16.4 Experimental Data 385
16.5 Current Issues 385
16.6 Practical Examples 387
16.6.1 Studies on Model Systems 387
16.6.2 FE Methods Applied to Pharmaceutically Relevant Systems 389
16.7 Miscellaneous Issues 395
16.8 Best Practices 396
16.9 Conclusions and Outlook 397
Acknowledgments 398
Abbreviations 398
References 399

17 Bioisosteres in De Novo Design 417
Nicholas C. Firth, Julian Blagg, and Nathan Brown
17.1 Introduction 417
17.2 History of Isosterism and Bioisosterism 418
17.3 Methods for Bioisosteric Replacement 421
17.3.1 Databases 422
17.3.1.1 BIOSTER 422
17.3.1.2 Cambridge Structure Database 422
17.3.1.3 ChEMBL 423
17.3.2 Descriptors 424
17.3.2.1 Physicochemical Properties 424
17.3.2.2 Molecular Topology 426
17.3.2.3 Molecular Shape 426
17.4 Exemplar Applications 427
17.4.1 Information-Based Bioisosteric Replacement 427
17.4.2 Drug Guru 429
17.4.3 SkelGen 431
17.5 Conclusions 433
Acknowledgments 433
References 434

18 Peptide Design by Nature-Inspired Algorithms 437
Jan A. Hiss and Gisbert Schneider
18.1 Template-Based Design 437
18.2 Nature-Inspired Optimization 441
18.2.1 Evolutionary Algorithms 444
18.2.2 Particle Swarm Optimization 446
18.2.3 Ant Colony Optimization 449
18.3 Worked Example: De Novo Design of MHC-I Binding Peptides
by Ant Colony Optimization 450
18.4 Chemical Modification 456
18.4.1 Backbone Cyclization 456
18.4.2 Stapling 458
18.4.3 End-Capping 458
18.4.4 Sugar-Coating 459
18.5 Conclusions and Outlook 460
19  *De Novo Computational Protein Design*  467

*Jeffery G. Saven*

19.1  Introduction  467
19.2  Elements of Computational Protein Design  470
19.2.1  Target Structures  470
19.2.2  Degrees of Freedom: Amino Acids and Side-Chain Conformations  470
19.2.2.1  Amino Acids  470
19.2.2.2  Side-Chain Conformations  471
19.2.3  Energy Functions  471
19.2.4  Solvation  472
19.2.5  Foldability Criteria and Negative Design  472
19.2.6  Sequence Search and Characterization  473
19.2.6.1  Monte Carlo  473
19.2.6.2  Dead-End Elimination  474
19.2.6.3  Mean Field Theory  475
19.2.6.4  Probabilistic Approach  475
19.3  Efforts in Theoretically Guided Protein Design  477
19.3.1  Toward Catalysis, Redox Activity, and Enzymes  477
19.3.2  *De Novo* Design and Redesign  478
19.3.3  Protein Reengineering  479
19.3.4  Cofactors and Nonbiological Protein Assemblies  480
19.3.5  Membrane Proteins  481
19.3.6  Protein–Protein Interactions and Protein Assemblies  483
19.4  Conclusion  485

Acknowledgments  485
References  486

20  *De Novo Design of Nucleic Acid Structures*  495

*Barbara Saccà, Andreas Sprengel, and Udo Feldkamp*

20.1  Introduction  495
20.2  DNA-Branched Structures  499
20.2.1  *De Novo* Design of DNA Junctions  499
20.2.2  Tile-to-Tile Binding  504
20.3  Scaffolded DNA Origami Design  505
20.3.1  Monolayer DNA Origami  506
20.3.1.1  Two-Dimensional Structures  506
20.3.1.2  Three-Dimensional Structures  509
20.3.2  Multilayer DNA Origami  509
20.4  Alternative DNA Designs: between Junctions and Origami  511
20.5 Conclusions 514
Acknowledgments 515
References 515

21 RNA Aptamer Design 519

Cindy Meyer, Ulrich Hahn, and Andrew E. Torda

21.1 Aptamers and Design 519
21.2 Riboswitches and Aptamers 520
21.3 SELEX 521
21.3.1 Introduction 521
21.3.2 The Method 522
21.3.3 Technical Challenges and Recent Developments in SELEX 526
21.4 Speeding Up SELEX by Computational Methods 526
21.4.1 Design of Structures 529
21.5 Structures and Probing Methods 530
21.6 Functional Analyses (In Vitro and In Vivo) 532
21.7 Problems 533
21.8 Future Perspectives 535
References 536

Index 543