

# CONTENTS

|                         |             |
|-------------------------|-------------|
| <i>Contributors</i>     | <i>xi</i>   |
| <i>Preface</i>          | <i>xvii</i> |
| <i>Volume in Series</i> | <i>xix</i>  |

## **1. Computational Design of Novel Protein Binders and Experimental Affinity Maturation** **1**

Timothy A. Whitehead, David Baker, and Sarel J. Fleishman

|                                                                          |    |
|--------------------------------------------------------------------------|----|
| 1. Introduction                                                          | 2  |
| 2. Computational Design of Binders Using Novel Scaffolds                 | 3  |
| 3. Target Selection                                                      | 4  |
| 4. Generating an Idealized Concept of the Hotspot                        | 5  |
| 5. Selecting Shape Complementary Scaffold Surfaces for Design            | 7  |
| 6. Interface Design                                                      | 8  |
| 7. Yeast Cell-Surface Display as a Screening Method for Designed Binders | 9  |
| 8. Affinity Maturation                                                   | 11 |
| 9. What Works, What Fails, and What It Means                             | 13 |
| Acknowledgments                                                          | 17 |
| References                                                               | 17 |

## **2. Mining Tertiary Structural Motifs for Assessment of Designability** **21**

Jian Zhang and Gevorg Grigoryan

|                              |    |
|------------------------------|----|
| 1. Introduction              | 22 |
| 2. MaDCaT                    | 25 |
| 3. Quantifying Designability | 32 |
| 4. Further Developments      | 36 |
| 5. Summary                   | 37 |
| Acknowledgments              | 37 |
| References                   | 37 |

## **3. Computational Methods for Controlling Binding Specificity** **41**

Oz Sharabi, Ariel Erijman, and Julia M. Shifman

|                                       |    |
|---------------------------------------|----|
| 1. Introduction                       | 42 |
| 2. Narrowing Down Binding Specificity | 44 |
| 3. Broadening Binding Specificity     | 52 |

|                                                                                                                                                                                                                                                                                    |            |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------|
| 4. Summary                                                                                                                                                                                                                                                                         | 58         |
| Acknowledgments                                                                                                                                                                                                                                                                    | 58         |
| References                                                                                                                                                                                                                                                                         | 58         |
| <b>4. Flexible Backbone Sampling Methods to Model and Design Protein Alternative Conformations</b>                                                                                                                                                                                 | <b>61</b>  |
| Noah Ollikainen, Colin A. Smith, James S. Fraser, and Tanja Kortemme                                                                                                                                                                                                               |            |
| 1. Introduction                                                                                                                                                                                                                                                                    | 62         |
| 2. Rosetta Moves to Model Alternative Conformations in X-Ray Density                                                                                                                                                                                                               | 64         |
| 3. Sequence Plasticity and Conformational Plasticity are Intertwined                                                                                                                                                                                                               | 76         |
| 4. Future Challenges                                                                                                                                                                                                                                                               | 82         |
| Acknowledgments                                                                                                                                                                                                                                                                    | 83         |
| References                                                                                                                                                                                                                                                                         | 83         |
| <b>5. OSPREY: Protein Design with Ensembles, Flexibility, and Provable Algorithms</b>                                                                                                                                                                                              | <b>87</b>  |
| Pablo Gainza, Kyle E. Roberts, Ivelin Georgiev, Ryan H. Lilien, Daniel A. Keedy, Cheng-Yu Chen, Faisal Reza, Amy C. Anderson, David C. Richardson, Jane S. Richardson, and Bruce R. Donald                                                                                         |            |
| 1. Introduction                                                                                                                                                                                                                                                                    | 88         |
| 2. OSPREY Design Principles                                                                                                                                                                                                                                                        | 89         |
| 3. Applications of OSPREY                                                                                                                                                                                                                                                          | 93         |
| 4. Protein Design in OSPREY                                                                                                                                                                                                                                                        | 94         |
| 5. Example: Predicting Drug Resistance Mutations Using OSPREY                                                                                                                                                                                                                      | 100        |
| 6. Future Directions and Availability                                                                                                                                                                                                                                              | 105        |
| Acknowledgments                                                                                                                                                                                                                                                                    | 105        |
| References                                                                                                                                                                                                                                                                         | 105        |
| <b>6. Scientific Benchmarks for Guiding Macromolecular Energy Function Improvement</b>                                                                                                                                                                                             | <b>109</b> |
| Andrew Leaver-Fay, Matthew J. O'Meara, Mike Tyka, Ron Jacak, Yifan Song, Elizabeth H. Kellogg, James Thompson, Ian W. Davis, Roland A. Pache, Sergey Lyskov, Jeffrey J. Gray, Tanja Kortemme, Jane S. Richardson, James J. Havranek, Jack Snoeyink, David Baker, and Brian Kuhlman |            |
| 1. Introduction                                                                                                                                                                                                                                                                    | 110        |
| 2. Energy Function Model                                                                                                                                                                                                                                                           | 112        |
| 3. Feature Analysis                                                                                                                                                                                                                                                                | 112        |
| 4. Maximum Likelihood Parameter Estimation with optE                                                                                                                                                                                                                               | 119        |
| 5. Large-Scale Benchmarks                                                                                                                                                                                                                                                          | 128        |

|                                                                                                                           |            |
|---------------------------------------------------------------------------------------------------------------------------|------------|
| 6. Three Proposed Changes to the Rosetta Energy Function                                                                  | 132        |
| 7. Conclusion                                                                                                             | 140        |
| Acknowledgments                                                                                                           | 140        |
| References                                                                                                                | 141        |
| <b>7. Molecular Dynamics Simulations for the Ranking, Evaluation, and Refinement of Computationally Designed Proteins</b> | <b>145</b> |
| Gert Kiss, Vijay S. Pande, and K.N. Houk                                                                                  |            |
| 1. Introduction                                                                                                           | 146        |
| 2. Inside-Out Computational Enzyme Design                                                                                 | 148        |
| 3. Filtering, Ranking, and Evaluation of Final Designs                                                                    | 149        |
| 4. Discerning Active from Inactive Designs with MD                                                                        | 151        |
| 5. MD Evaluation Examples                                                                                                 | 156        |
| 6. MD Refinement Examples                                                                                                 | 161        |
| 7. Molecular Dynamics Simulations: Preparation and Setup                                                                  | 166        |
| 8. Conclusions                                                                                                            | 167        |
| Acknowledgments                                                                                                           | 168        |
| References                                                                                                                | 168        |
| <b>8. Multistate Protein Design Using CLEVER and CLASSY</b>                                                               | <b>171</b> |
| Christopher Negron and Amy E. Keating                                                                                     |            |
| 1. Introduction: Accomplishments and Limitations of Structure-Based Design                                                | 172        |
| 2. Theory                                                                                                                 | 173        |
| 3. Benefits Offered by Cluster Expansion in Protein Modeling and Design                                                   | 175        |
| 4. How to Run a Cluster Expansion with CLEVER 1.0                                                                         | 178        |
| 5. GenSeqs                                                                                                                | 178        |
| 6. CEFTrFILE                                                                                                              | 179        |
| 7. CEEnergy                                                                                                               | 180        |
| 8. Cluster Expansion Case Study                                                                                           | 180        |
| 9. Using Cluster Expansion with Integer Linear Programming                                                                | 183        |
| 10. CLASSY Applied to Multistate Design                                                                                   | 185        |
| 11. Conclusion                                                                                                            | 187        |
| Acknowledgments                                                                                                           | 188        |
| References                                                                                                                | 188        |
| <b>9. Using Analyses of Amino Acid Coevolution to Understand Protein Structure and Function</b>                           | <b>191</b> |
| Orr Ashenberg and Michael T. Laub                                                                                         |            |
| 1. Introduction                                                                                                           | 192        |
| 2. Predicting Specificity Determining Residues Using MI                                                                   | 199        |

|                                                                                                               |            |
|---------------------------------------------------------------------------------------------------------------|------------|
| 3. Concluding Remarks                                                                                         | 209        |
| Acknowledgments                                                                                               | 209        |
| References                                                                                                    | 209        |
| <b>10. Evolution-Based Design of Proteins</b>                                                                 | <b>213</b> |
| Kimberly A. Reynolds, William P. Russ, Michael Socolich, and Rama Ranganathan                                 |            |
| 1. Introduction                                                                                               | 214        |
| 2. SCA: The Pattern of Evolutionary Constraint in Proteins                                                    | 215        |
| 3. SCA-Based Protein Design                                                                                   | 222        |
| 4. SCA-Based Parsing of Protein Stability and Function                                                        | 227        |
| 5. Future Monte Carlo Strategies for Exploring Sequence Space                                                 | 230        |
| 6. Conclusion                                                                                                 | 234        |
| Acknowledgments                                                                                               | 234        |
| References                                                                                                    | 234        |
| <b>11. Protein Engineering and Stabilization from Sequence Statistics: Variation and Covariation Analysis</b> | <b>237</b> |
| Venuka Durani and Thomas J. Magliery                                                                          |            |
| 1. Introduction                                                                                               | 238        |
| 2. Case Study: BPTI                                                                                           | 239        |
| 3. Acquiring an MSA                                                                                           | 240        |
| 4. Relative Entropies: Quantifying the Degree of Positional Variation                                         | 244        |
| 5. Mutual Information: Quantifying the Degree of Covariation                                                  | 246        |
| 6. Protocol for Predicting Stabilizing Mutations                                                              | 253        |
| 7. Summary                                                                                                    | 253        |
| References                                                                                                    | 254        |
| <b>12. Enzyme Engineering by Targeted Libraries</b>                                                           | <b>257</b> |
| Moshe Goldsmith and Dan S. Tawfik                                                                             |            |
| 1. Introduction                                                                                               | 258        |
| 2. Screening Versus Selection                                                                                 | 259        |
| 3. The Merits of a Direct Screen: The Nerve Agent-Detoxifying Enzymes                                         | 261        |
| 4. Hedging the Bets: Mutational Spiking Approaches                                                            | 264        |
| 5. Rational and Analytical Library Designs                                                                    | 266        |
| 6. Summary                                                                                                    | 278        |
| References                                                                                                    | 278        |

|                                                                                                                      |            |
|----------------------------------------------------------------------------------------------------------------------|------------|
| <b>13. Generation of High-Performance Binding Proteins for Peptide Motifs by Affinity Clamping</b>                   | <b>285</b> |
| Shohei Koide and Jin Huang                                                                                           |            |
| 1. Introduction                                                                                                      | 286        |
| 2. The Affinity Clamping Concept                                                                                     | 287        |
| 3. Design of Affinity Clamps                                                                                         | 289        |
| 4. Production and Characterization of Affinity Clamps                                                                | 298        |
| 5. Applications of Affinity Clamps                                                                                   | 299        |
| 6. Conclusion                                                                                                        | 300        |
| Acknowledgments                                                                                                      | 300        |
| References                                                                                                           | 300        |
| <b>14. Engineering Fibronectin-Based Binding Proteins by Yeast Surface Display</b>                                   | <b>303</b> |
| Tiffany F. Chen, Seymour de Picciotto, Benjamin J. Hackel, and K. Dane Wittrup                                       |            |
| 1. Introduction                                                                                                      | 304        |
| 2. Engineering and Screening Approach of Fn3s                                                                        | 307        |
| 3. Analysis of Individual Clones                                                                                     | 322        |
| 4. Summary                                                                                                           | 324        |
| Acknowledgments                                                                                                      | 325        |
| References                                                                                                           | 325        |
| <b>15. Engineering and Analysis of Peptide-Recognition Domain Specificities by Phage Display and Deep Sequencing</b> | <b>327</b> |
| Megan E. McLaughlin and Sachdev S. Sidhu                                                                             |            |
| 1. Introduction                                                                                                      | 328        |
| 2. Directed Evolution of PDZ Variants                                                                                | 329        |
| 3. Peptide Profiling of PDZ Variants                                                                                 | 341        |
| 4. Summary                                                                                                           | 348        |
| Acknowledgment                                                                                                       | 348        |
| References                                                                                                           | 348        |
| <b>16. Efficient Sampling of SCHEMA Chimera Families to Identify Useful Sequence Elements</b>                        | <b>351</b> |
| Pete Heinzelman, Philip A. Romero, and Frances H. Arnold                                                             |            |
| 1. Introduction                                                                                                      | 352        |
| 2. SCHEMA Chimera Family Design Overview                                                                             | 352        |



|                                                                                                           |            |
|-----------------------------------------------------------------------------------------------------------|------------|
| 3. Prediction of Thermostable Chimeras by Linear Regression Modeling                                      | 355        |
| 4. Summary                                                                                                | 367        |
| Acknowledgments                                                                                           | 367        |
| References                                                                                                | 368        |
| <b>17. Protein Switch Engineering by Domain Insertion</b>                                                 | <b>369</b> |
| Manu Kanwar, R. Clay Wright, Amol Date, Jennifer Tullman, and Marc Ostermeier                             |            |
| 1. Introduction                                                                                           | 370        |
| 2. Creation of Random Double-Stranded Breaks in Plasmids Containing the Acceptor DNA                      | 373        |
| 3. Repair, Purification and Dephosphorylation of Acceptor DNA                                             | 379        |
| 4. Preparation of Insert DNA                                                                              | 381        |
| 5. Ligation, Transformation, Recovery, and Storage of the Library                                         | 385        |
| 6. Characterization of the Library                                                                        | 386        |
| Acknowledgments                                                                                           | 387        |
| References                                                                                                | 387        |
| <b>18. Design of Chimeric Proteins by Combination of Subdomain-Sized Fragments</b>                        | <b>389</b> |
| José Arcadio Fariás Rico and Birte Höcker                                                                 |            |
| 1. Introduction                                                                                           | 390        |
| 2. Selecting the Starting Structures for Chimera Design                                                   | 395        |
| 3. Evaluation and Optimization of the Chimera                                                             | 398        |
| 4. Summary and Final Considerations                                                                       | 404        |
| References                                                                                                | 404        |
| <b>19. <math>\alpha</math>-Helix Mimicry with <math>\alpha/\beta</math>-Peptides</b>                      | <b>407</b> |
| Lisa M. Johnson and Samuel H. Gellman                                                                     |            |
| 1. Introduction                                                                                           | 408        |
| 2. Helical Secondary Structures from $\beta$ -Peptides and $\alpha/\beta$ -Peptides                       | 409        |
| 3. Biological Function from Helical $\beta$ -Peptides                                                     | 411        |
| 4. $\alpha$ -Helix Mimicry with $\alpha/\beta$ -Peptides                                                  | 414        |
| 5. Toward a General Approach for $\alpha$ -Helix Mimicry with Protease-Resistant $\alpha/\beta$ -Peptides | 423        |
| Acknowledgments                                                                                           | 425        |
| References                                                                                                | 425        |
| <b>Author Index</b>                                                                                       | <b>431</b> |
| <b>Subject Index</b>                                                                                      | <b>453</b> |