

*Hans-Dieter Höltje, Wolfgang Sippl, Didier Rognan,
and Gerd Folkers*

Molecular Modeling

Basic Principles and Applications

Third, Revised and Expanded Edition



**WILEY-
VCH**

WILEY-VCH Verlag GmbH & Co. KGaA

Contents

Preface to the Third Edition X

| | | |
|----------|---|----------|
| 1 | Introduction | 1 |
| 1.1 | Modern History of Molecular Modeling | 2 |
| 1.2 | Do Today's Molecular Modeling Methods Only Make Pictures of the Lukretian World or Do They Make Anything More? | 3 |
| 1.3 | What are Models Used For? | 4 |
| 1.4 | Molecular Modeling Uses all Four Kinds for Model Building | 5 |
| 1.5 | The Final Step Is <i>Design</i> | 5 |
| 1.6 | Scope of the Book | 6 |
| | | |
| 2 | Small Molecules | 9 |
| 2.1 | Generation of 3D Coordinates | 9 |
| 2.1.1 | Crystal Data | 9 |
| 2.1.2 | Fragment Libraries | 10 |
| 2.1.3 | Conversion of 2D Structural Data into 3D Form | 12 |
| | References | 15 |
| 2.2 | Computational Tools for Geometry Optimization | 16 |
| 2.2.1 | Force Fields | 16 |
| 2.2.2 | Geometry Optimization | 19 |
| 2.2.3 | Energy-minimizing Procedures | 21 |
| 2.2.4 | Use of Charges, Solvation Effects | 23 |
| 2.2.5 | Quantum Mechanical Methods | 24 |
| | References | 29 |
| 2.3 | Conformational Analysis | 32 |
| 2.3.1 | Conformational Analysis Using Systematic Search Procedures | 34 |
| 2.3.2 | Conformational Analysis Using Monte Carlo Methods | 37 |
| 2.3.3 | Conformational Analysis Using Molecular Dynamics | 39 |
| 2.3.4 | Which Is the Method of Choice? | 44 |
| | References | 46 |

| | | |
|----------|--|------------|
| 2.4 | Determination of Molecular Interaction Potentials | 50 |
| 2.4.1 | Molecular Electrostatic Potentials (MEPs) | 50 |
| 2.4.2 | Molecular Interaction Fields | 57 |
| 2.4.3 | Display of Properties on a Molecular Surface | 66 |
| | References | 66 |
| | Further Reading | 69 |
| 2.5 | Pharmacophore Identification | 70 |
| 2.5.1 | Molecules to be Matched | 70 |
| 2.5.2 | Atom-by-atom Superposition | 72 |
| 2.5.3 | Superposition of Molecular Fields | 74 |
| | References | 75 |
| 2.6 | 3D QSAR Methods | 77 |
| 2.6.1 | The CoMFA Method | 77 |
| 2.6.2 | Other CoMFA-related Methods | 81 |
| 2.6.3 | More 3D QSAR Methods | 83 |
| 2.6.4 | Receptor-based 3D QSAR | 84 |
| 2.6.5 | Reliability of 3D QSAR Models | 86 |
| | References | 87 |
| | Further Reading | 91 |
| 3 | A Case Study for Small Molecule Modeling: Dopamine D₃ Receptor Antagonists | 93 |
| 3.1 | A Pharmacophore Model for Dopamine D ₃ Receptor Antagonists | 93 |
| 3.1.1 | The Aromatic–Basic Fragment | 99 |
| 3.1.2 | The Spacer | 100 |
| 3.1.3 | The Aromatic–Amidic Residue | 101 |
| 3.1.4 | Resulting Pharmacophore | 102 |
| 3.1.5 | Molecular Interaction Fields | 102 |
| 3.2 | 3D QSAR Analysis | 104 |
| 3.2.1 | Variable Reduction and PLS Model | 104 |
| 3.2.2 | Validation of the Model | 107 |
| 3.2.3 | Prediction of External Ligands | 108 |
| | References | 110 |
| 4 | Introduction to Comparative Protein Modeling | 111 |
| 4.1 | Where and How to Get Information on Proteins | 111 |
| | References | 115 |
| 4.2 | Terminology and Principles of Protein Structure | 116 |
| 4.2.1 | Conformational Properties of Proteins | 116 |
| 4.2.2 | Types of Secondary Structural Elements | 119 |
| 4.2.3 | Homologous Proteins | 122 |
| | References | 124 |

| | | |
|----------|---|------------|
| 4.3 | Comparative Protein Modeling | 126 |
| 4.3.1 | Procedures for Sequence Alignments | 127 |
| 4.3.2 | Determination and Generation of Structurally Conserved Regions (SCRs) | 133 |
| 4.3.3 | Construction of Structurally Variable Regions (SVRs) | 135 |
| 4.3.4 | Side-Chain Modeling | 136 |
| 4.3.5 | Distance Geometry Approach | 138 |
| 4.3.6 | Secondary Structure Prediction | 139 |
| 4.3.7 | Threading Methods | 141 |
| | References | 144 |
| 4.4 | Optimization Procedures – Model Refinement – Molecular Dynamics | 149 |
| 4.4.1 | Force Fields for Protein Modeling | 149 |
| 4.4.2 | Geometry Optimization | 150 |
| 4.4.3 | The Use of Molecular Dynamics Simulations in Model Refinement | 151 |
| 4.4.4 | Treatment of Solvated Systems | 153 |
| 4.4.5 | Ligand-binding Site Complexes | 155 |
| | References | 155 |
| 4.5 | Validation of Protein Models | 158 |
| 4.5.1 | Stereochemical Accuracy | 158 |
| 4.5.2 | Packing Quality | 164 |
| 4.5.3 | Folding Reliability | 166 |
| | References | 169 |
| 4.6 | Properties of Proteins | 173 |
| 4.6.1 | Electrostatic Potential | 173 |
| 4.6.2 | Interaction Potentials | 177 |
| 4.6.3 | Hydrophobicity | 177 |
| | References | 178 |
| 5 | Virtual Screening and Docking | 181 |
| 5.1 | Preparation of the Partners | 181 |
| 5.1.1 | Preparation of the Compound Library | 181 |
| 5.1.2 | Representation of Proteins and Ligands | 186 |
| 5.2 | Docking Algorithms | 189 |
| 5.2.1 | Incremental Construction Methods | 189 |
| 5.2.2 | Genetic Algorithms | 191 |
| 5.2.3 | Tabu Search | 192 |
| 5.2.4 | Simulated Annealing and Monte Carlo Simulations | 194 |
| 5.2.5 | Shape-fitting Methods | 195 |
| 5.2.6 | Miscellaneous Approaches | 195 |
| 5.3 | Scoring Functions | 196 |
| 5.3.1 | Empirical Scoring Functions | 196 |

| | | |
|-------|--|-----|
| 5.3.2 | Force-field-based Scoring Functions | 198 |
| 5.3.3 | Knowledge-based Scoring Functions | 198 |
| 5.3.4 | Critical Overview of Fast Scoring Functions | 199 |
| 5.4 | Postfiltering Virtual Screening Results | 200 |
| 5.4.1 | Filtering by Topological Properties | 200 |
| 5.4.2 | Filtering by Consensus Mining Approaches | 200 |
| 5.4.3 | Filtering by Combining Computational Procedures | 201 |
| 5.4.4 | Filtering by Chemical Diversity | 201 |
| 5.4.5 | Filtering by Visual Inspection | 202 |
| 5.5 | Comparison of Different Docking and Scoring Methods | 202 |
| 5.6 | Examples of Successful Virtual Screening Studies | 203 |
| 5.7 | Outlook | 206 |
| | References | 207 |
| | | |
| 6 | Scope and Limits of Molecular Docking | 217 |
| 6.1 | Docking in the Polar Active Site that Contains Water Molecules | 218 |
| 6.2 | Including Cofactor in Docking? | 225 |
| 6.3 | Impact of Tautomerism on Docking | 227 |
| | References | 229 |
| | Further Reading | 231 |
| | | |
| 7 | Chemogenomic Approaches to Rational Drug Design | 233 |
| 7.1 | Description of Ligand and Target Spaces | 235 |
| 7.1.1 | Ligand Space | 236 |
| 7.1.2 | Target Space | 238 |
| 7.1.3 | Protein–Ligand Space | 240 |
| 7.2 | Ligand-based Chemogenomic Approaches | 242 |
| 7.2.1 | Annotating Ligand Libraries | 242 |
| 7.2.2 | Privileged Structures | 244 |
| 7.2.3 | Ligand-based <i>In silico</i> Screening | 246 |
| 7.3 | Target-based Chemogenomic Approaches | 249 |
| 7.3.1 | Sequence-based Comparisons | 249 |
| 7.3.2 | Structure-based Comparisons | 251 |
| 7.4 | Target-Ligand-based Chemogenomic Approaches | 254 |
| 7.4.1 | Chemical Annotation of Target Binding Sites | 254 |
| 7.4.2 | Two-dimensional Searches | 256 |
| 7.4.3 | Three-dimensional Searches | 256 |
| 7.5 | Concluding Remarks | 258 |
| | References | 258 |

| | | |
|----------|--|------------|
| 8 | A Case Study for Protein Modeling: the Nuclear Hormone Receptor CAR as an Example for Comparative Modeling and the Analysis of Protein-Ligand Complexes | 265 |
| 8.1 | The Biochemical and Pharmacological Description of the Problem | 265 |
| 8.1.1 | Nuclear Hormone Receptor Superfamily | 265 |
| 8.1.2 | Molecular Architecture and Activation Mechanisms of Nuclear Hormone Receptors | 265 |
| 8.1.3 | The Human Constitutive Active Androstan Receptor (CAR) | 267 |
| 8.1.4 | CAR Ligands | 267 |
| 8.2 | Comparative Modeling of the Human Nuclear Hormone Receptor CAR | 268 |
| 8.2.1 | Choosing Appropriate Template Structures | 269 |
| 8.2.2 | Homology Modeling of the Human CAR | 271 |
| 8.2.3 | Setting up the System for the Molecular Dynamics Simulations | 271 |
| 8.3 | Analysis of the Models that Emerged from MD Simulations | 272 |
| 8.3.1 | Atomic Fluctuations | 272 |
| 8.3.2 | AF-2 Interaction Domain | 275 |
| 8.3.3 | Deciphering the Structural Basis for Constitutive Activity of Human CAR | 276 |
| 8.3.4 | Coactivator Binding | 278 |
| 8.4 | Analysis of CAR Mutants | 279 |
| 8.4.1 | Identifying Important Amino Acids for CAR Activation | 279 |
| 8.4.2 | MD Simulations of Selected CAR Mutants | 282 |
| 8.5 | Modeling of CAR-Ligand Complexes | 284 |
| 8.6 | The CAR X-ray Structure Comes into Play | 286 |
| 8.6.1 | How Accurate is the Generated CAR Model? | 286 |
| 8.6.2 | Docking Studies Using the CAR X-ray Structure | 288 |
| 8.6.3 | The Basis for Constitutive Activity Revisited | 289 |
| 8.7 | Virtual Screening for Novel CAR Activators | 292 |
| 8.8 | Concluding Remarks | 295 |
| | References | 296 |
| | Index | 299 |