

Preface

This book is a collection of the lecture notes for a course on computational structural biology taught at the Program on Bioinformatics and Computational Biology at Iowa State University, USA. It is, in a certain sense, a summary of what I have learned and understood about computational structural biology based on my own research experience, and is more of a review from the perspective of a computer scientist or applied mathematician on the subject.

The book is focused on introducing computational and mathematical problems, their solution methods, and the research developments. Efforts have been made to keep the contents self-contained and not too technical, so that readers not previously exposed to related subjects can still understand the materials and appreciate the biological importance and mathematical novelty of the research in the field.

The field of computational structural biology is new and still developing. As such, it is hard to define the exact scope of the field, as different groups of researchers may prefer different names with different research goals and scopes in their minds, such as structural bioinformatics, structural genomics, or biomolecular modeling, to name a few. Here, we just choose the current name, not exclusively, and restrict our discussion to the use of computational methods for the study of the structures and dynamics of biomolecules and, in particular, proteins.

Structures and their dynamic properties are key to the understanding of the biological functions of proteins. However, in general, they are difficult to determine. Both experimental and theoretical approaches have been used. They all require intensive mathematical

computing, including data analysis, model building, structure prediction, and dynamics simulation. Mathematical problems arise in various stages of investigations on structure and dynamics. Some of them can be resolved effectively by using existing mathematical tools, while others present remarkable computational challenges. The development of computational structural biology thus follows the great need for mathematics and computation in structural research.

In Chap. 1, we discuss the importance of structures for the understanding of proteins and related biological functions. Two classes of computational problems, structure determination and dynamics simulation, as the main subjects of computational structural biology, are introduced. A review on the famous problem of protein folding is given.

The topics related to X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy, two major experimental approaches for structure determination, are covered in Chaps. 2 and 3, respectively, with emphasis on the phase problem in X-ray crystallography and the distance geometry problem in NMR and their solutions. Two important computational methods, potential energy minimization and molecular dynamics simulation, as well as their extensive applications to protein structures and dynamics are discussed in Chaps. 4 and 5.

The knowledge-based approach to structure prediction, an active area of research in recent years, is reviewed in Chap. 6, with a great deal of details left for readers to explore further. The development of other related fields, including computational genomics and systems biology, and their connection with computational structural biology are also discussed.

To be self-contained, some general knowledge on algorithms and numerical methods are given with relatively rigorous mathematical descriptions in Appendices A and B. In contrast, in the main chapters of the book, the presentation is made purposely without formal mathematical statements (such as theorems and proofs). The bibliography is provided in the form of further readings attached to each chapter and grouped under specific topics. They are selected readings, not complete references, and are also given in certain suggested reading orders.

I would like to thank my students and colleagues, Ajith Gunaratne, Feng Cui, Qunfeng Dong, Rahul Ravindrudu, Vladimir Sukhoy, Peter Vedell, Atilla Sit, Xiaoyong Sun, Di Wu, and Wonbin Young, for sharing their knowledge and providing valuable insights while working with me. The constant discussions among us on research in computational biology have been the major source of motivation for the writing of the book. I would also like to thank my biology colleagues David Shalloway, George Phillips, and Robert Jernigan for working with me and introducing me to the fascinating field of computational biology; and my optimization advisers and colleagues John Dennis, Robert Bixby, Tom Coleman, Jorge Moré, Richard Tapia, Yin Zhang, Guangye Li, Yinyu Ye, and Yaxiang Yuan for teaching me optimization and encouraging me to explore scientific applications. I would like to acknowledge the strong support provided by the Department of Mathematics, the Graduate Program on Bioinformatics and Computational Biology, and the Baker Center for Bioinformatics and Biological Statistics, Iowa State University. I have also benefited from the support provided by the Mathematical Biosciences Institute at Ohio State University, USA, the Institute for Mathematics and Its Applications at the University of Minnesota, USA, and the Institute of Computational Mathematics at the Chinese Academy of Sciences, China, for allowing me to visit the institutes while finishing the book. Finally, I would like to thank Peng Wu for reading the initial draft of the book and providing helpful suggestions; my friend and colleague Dr Chunguang Sun at World Scientific Publishing Company (WSPC) for suggesting the book project; and the WSPC staff, Wanda Tan, Azouri Bongso, and Mei-Lian Ho, for their professional help and friendly cooperation on text organization, cover design, and proof-reading.

Z. Wu