

Preface

The past decade has witnessed not only a flood of protein sequence and structure data generated by large-scale genomic sequencing and structural genomics projects, but also an ensuing growth of size and number of databases and computer programs designed to manage and process these data. The multitude of bioinformatic tools available to molecular biologists offers multiple solutions to various steps of process sequence-structure-function analyses. Often the choice of which tool to use depends more on its popularity among relatively naïve *users*, sometimes stemming from the availability of an intuitive web-server interface, rather than on an understanding of the underlying principles or on the user's ability to utilize all the information returned by the program, including the assessment of confidence of the results. Being educated and trained in molecular biology and biochemistry and self-taught in bioinformatics, I am interested in both the development of computational tools and their optimal application in the realm of experimental biology, especially in the studies of protein-nucleic acid interactions. Despite the abundance of literature on bioinformatics and on molecular biology of proteins that interact with nucleic acids, there are few (if any) timely volumes dedicated to the synthesis of these two research areas. Hence, I was delighted to accept the invitation to act as an editor of a "*Practical Bioinformatics*" volume of *Nucleic Acids and Molecular Biology* and to consolidate key bioinformatic methods for studying protein sequence-structure-function relationships into a convenient source.

This volume is mainly for the biochemist or molecular biologist who wants to analyze, search or manipulate protein structure or sequence data and to integrate these analyses with their experimental investigations to interpret the obtained results or to plan further studies better. Thus, the first part of the volume comprises reviews of methodology solicited from developers of bioinformatic software (with the emphasis on methods that explicitly utilize experimental information and/or are designed to guide experimental research), while the second part comprises useful strategies for studying protein function with the aid of bioinformatics, described in the form of "case

studies” by at-the-bench scientists. Methods and strategies range from protein structure prediction by template-dependent (comparative modeling, fold-recognition) and template-independent (*ab initio*) approaches, to prediction of protein–protein and protein–nucleic acid interactions, to identification of proteins exerting a defined function or prediction of the function for newly identified proteins. In the spirit of this series, all case studies involve analyses of proteins involved in interactions with nucleic acids – from ribosome assembly and structure, to posttranscriptional RNA modification, to DNA restriction and repair.

The bioinformatics field is a very fast-moving one, and every effort was made to produce this volume as rapidly as possible so the methods would be timely. In this regard, I am grateful to all the authors for taking their time to contribute and for adhering to a set of rigid deadlines; without their participation this volume would not have been possible. I hope that *Practical Bioinformatics* will serve as a useful compendium of methods both to newcomers in the field of bioinformatics-aided experimental molecular biology and biochemistry as well as to scientists actively engaged in research in this area.

Warsaw, July 2003

Janusz M. Bujnicki