Preface to the Second Edition

Welcome to the second edition! Since the publication of the first edition, the research area of protein structural informatics has continued to grow in volume and significance. A search of PubMed for 'protein structural bioinformatics' shows around 1000 papers in 2009 when the first edition was published, doubling to over 2000 in 2015. In the same period, the Protein Data Bank has similarly almost doubled, breaching 100,000 entries in 2014. Nevertheless, the gap between the protein sequences and structures continues to grow, as new technologies allow cheap and facile sequencing of previously intractable organisms and even of entire environments. Protein structural bioinformatics offers a computational route to bridge this gap by predicting structures for uncharacterised families. Those structures can then be analysed by a wide variety of further bioinformatics algorithms to shed light on their function. These two interlinking research areas are the topic of this book.

This second edition contains three chapters addressing areas not covered in the first edition. Each is contributed by world-leading experts in the field. The remaining chapters are all revised, many dramatically, to reflect seven years of fast-moving bioinformatics research with one chapter being entirely replaced. As previously, there are two sections covering first methods to generate or infer structure and secondly structure-based function annotation. Naturally, such a division is never clear-cut as prediction of a structure may simultaneously inform about its likely functions. For example, annotation of an intrinsically disordered region would immediately suggest, in eukaryotes at least, a role in protein-protein interaction since such stretches frequently harbour linear motifs bound by recognition modules on partner proteins.

The first new chapter, Chap. 2, covers arguably the most exciting development in protein bioinformatics of recent years, namely the new-found ability to accurately predict contacting residue pairs through covariance analysis of large multiple sequence alignments. These contact predictions have a wide and still expanding range of applications. Most obviously, the data allow for protein structure prediction in conjunction either with protein distance geometry methods or, more effectively, by synergistic incorporation into fragment assembly ab initio modelling

methods. The contact predictions also inform on the likely harmfulness of single amino acid polymorphisms (SAPs) and allow for better prediction of protein-protein interactions. Prediction of protein-protein complex structures, both between globular domains and between a domain and a short linear motif, is the subject of the new Chap. 8. A full accounting of protein-protein interactions in cells is crucial for the future prospects of integrative systems-level methods, while structural knowledge of interfaces again contributes to prediction of the consequences of SAPs. The third new arrival, Chap. 7, covers predictions of amyloid structure in proteins. Such structure is of huge biomedical interest, underlying diseases such as Parkinson's and Alzheimer's, but is equally intriguing for the normal physiological roles of 'functional amyloids'. Finally, the new Chap. 10 text covers the fascinating variety of means by which structural bioinformatics can mark up a structure, experimental or modelled, for likely functional pockets and patches on the protein surface.

The methods covered in this book comprise a comprehensive toolkit to address future challenges in protein structure, function and evolution. Recent papers open up new viewpoints on protein evolution (Alva et al. 2015; Edwards and Deane 2015) and on the amenability of different folds to functional innovation (Toth-Petroczy and Tawfik 2014), treat the biophysical consequences of protein ageing (de Graff et al. 2016) and even reveal oversights in our accounting of molecular interactions (Newberry and Raines 2016). Clearly, exciting times lie ahead for protein bioinformaticians!

Liverpool, UK Daniel J. Rigden

References

Alva V, Soding J, Lupas AN (2015) A vocabulary of ancient peptides at the origin of folded proteins. Elife 4:e09410

de Graff AM, Hazoglou MJ, Dill KA (2016) Highly charged proteins: the achilles' heel of aging proteomes. Structure 24(2):329–336

Edwards H, Deane CM (2015) Structural bridges through fold space. PLoS Comput Biol 11(9): e1004466

Newberry RW, Raines RT (2016) A prevalent intraresidue hydrogen bond stabilizes proteins. Nat Chem Biol 12(12):1084–1088

Toth-Petroczy A, Tawfik DS (2014) The robustness and innovability of protein folds. Curr Opin Struct Biol 26:131–138

http://www.springer.com/978-94-024-1067-9

From Protein Structure to Function with Bioinformatics

J. Rigden, D. (Ed.)
2017, XV, 503 p. 86 illus., 75 illus. in color., Hardcover
ISBN: 978-94-024-1067-9