

Book Review

Even trickier than origami: challenges and advances in our understanding of protein folding

A Oberst

Cell Death and Differentiation Rome Editorial Office; Correspondence to: drew_oberst@hotmail.com

From Protein Folding to New Enzymes. Biochemical Society Symposium 68. Edited by A Berry and SE Radford. Portland Press, London: 2001. Pp 156. ISBN: 1-85578-143-3

The question of how a newly-synthesized polypeptide chain folds to form a functional protein is one of the most intriguing – and complex – in structural biology. The Annual Symposium of the Biochemical Society, held at the University of Leeds, UK, in April 2000, focused on the recent advances in our understanding of the folding and function of proteins. The book 'From Protein Folding to New Enzymes,' edited by A Berry and SE Radford, contains a representative collection of studies presented at this conference, and provides an informative summary of recent advances and future potential in the fast-evolving field of protein folding and structure.

The study of protein folding is a tricky business, and this book addresses many of the pitfalls – as well as their solutions – inherent to the field. *In vivo*, nascent polypeptide chains begin folding as they emerge from the ribosome, and their folding necessarily takes place in the complex environment of the living cell. The process of forming a functional protein from an emerging amino acid chain is therefore subject to the influence, both positive and negative, of a wide variety of pre-existing molecules, including other nascent polypeptides and a variety of protein chaperones. Further complicating the scene, the folding and function of many proteins is significantly influenced by the specific cellular environments in which they operate, the most obvious example being membrane proteins. Such a high degree of complexity is impossible to duplicate with *in vitro* work; such studies generally analyze the folding of complete amino acid chains from a fully-denatured state, in the absence of the influence of other proteins, and in the artificial chemical environment selected by the investigator. Despite the significant environmental differences, most proteins can be coaxed to fold into their functional states *in vitro* by careful choice of conditions; recent developments in high-speed microprocessing have also led to computer modeling of protein folding. Thought both *in vitro* and computer simulation studies provide a clearer view of the folding process undergone by specific proteins, they don't necessarily reflect the processes taking place in a living cell. As so often happens in biology, researchers are left to reconcile 'clean-but-artificial' *in vitro* data with 'dirty-but-real' *in vivo* data.

This book discusses many recent developments in the struggle to understand the highly complex interactions that govern protein folding both *in vivo* (including the function of protein chaperones) and *in vitro* (such as reproduction of

the membrane protein folding environment). It also explores the effects and mechanisms of protein misfolding and aggregation, the interplay between simulation and experiment in such studies, and the potential application of our knowledge of protein folding and function the emerging field of protein engineering.

The discussion of protein folding starts off with a bang, in the form of an extremely interesting paper by Christopher Dobson entitled 'Protein Folding and its links to Human Disease'. This paper presents an excellent introduction to methods currently used to study and visualize the folding process, and explores what happens when proteins misfold. The fundamental facet of protein folding diseases presented by Dobson is that of aggregation, which is blamed for both Alzheimer's and Creutzfeldt-Jakob disease. Protein aggregation occurs when mutation destabilizes the native-state protein, increasing the occurrence of the less-stable folding intermediates. The partially-folded nature of these intermediates presents residues that would normally be sequestered inside the native proteins; these in turn can lead to 'self-association' with other copies of the same destabilized protein, leading eventually to the formation of long, regular 'amyloid fibrils' composed of partially unfolded, aggregated proteins. Once fibril formation is initiated, the partially-formed fibril acts as a template for further aggregation, leading to accelerated fibril growth and indeed explaining the infectious nature of the prion diseases (a concept further discussed in a later paper by Serio *et al*). Dobson also presents the startling finding that, while the currently known amyloid diseases are caused by the aggregation of a select few proteins, aggregation is an inherent property of all polypeptides. Dobson explains that amyloid fibrils are made up of repeating β -sheet-like structure, in which stabilizing interactions are formed primarily by the peptide backbone; conditions have been found under which a wide variety of proteins not associated with disease, including proteins with little or no native β -sheet structure, aggregate into fibrils. In light of the finding that all polypeptides have the potential to form disease-causing aggregates, the rarity of such diseases can be viewed as a testament to the amazing efficiency with which proteins fold, and with which misfolded proteins are disposed of, in the *in vivo* folding system.

Some of the key components of this system are addressed by a pair of papers on the molecular chaperones

involved in protein folding. As already mentioned, semi-folded proteins have a tendency toward aggregation, often with disastrous consequences for the cell and, more generally, the organism. This problem is further aggravated by the high concentration of macromolecules and the close proximity of nascent polypeptides emerging from the ribosome. While most proteins can fold to their functional states without help in the sterile environment of *in vitro* studies, some intervention is needed to prevent aggregation of partially-folded proteins in the chaotic environment of the living cell. This intervention comes in the form of the molecular chaperone systems, which are summarized in the context of *E.coli* in a paper by Naylor and Hartl. The GroEL-GroES system functions by providing nascent proteins with a sanctuary from the hectic cellular environment. GroEL forms 'boxes' that, upon binding of seven molecules of ATP, open to allow a partially-folded protein substrate of up to 60 kDa to enter; GroES forms the lid to this box, which allows nascent proteins to fold in a sheltered environment. The second molecular chaperone system, the DnaK/DnaJ/GrpE system, promotes proper folding by shielding hydrophobic sections of nascent polypeptides against aggregation in an ATP-dependent manner. The structure and substrate specificity of the members of this system are discussed in detail, as is the reaction cycle by which the system works. Further exploration of the key protein component of this system, DnaK, is provided in a paper by Swain and coworkers, which details the effort to define the structure of the substrate-free DnaK molecule, an investigation which has yielded further information about the environmental sensitivity of this molecule's substrate binding properties.

Another pair of papers deals with recent developments in the folding of the most abundant protein secondary structure, the α -helix. A paper by Doig and co-workers details the latest findings in the effort to refine the helix-coil theoretical model, which describes the folding of the α -helix from a random coil. These findings focus on the incorporation of factors for the unique positions at the N and C caps of the helix as well as specific side chain-side chain interactions, and also surveys common helix lengths in the protein database. Close study of side chain-side chain interactions has led the authors of this paper to propose a 'new' view of protein folding energetics, which takes into account rotamer strain. The 'old' view holds that folded proteins are stabilized by the hydrophobic effect and hydrogen bonding, and destabilized by the loss of conformational entropy inherent in the folding event. The authors' studies of side chain conformation in α -helical amino acids indicate that the contribution of rotameric strain, which arises when side chain dihedral angles are shifted from their most stable conformation during folding, also contributes significantly to protein destabilization. The authors therefore propose that four factors, all of similar magnitude, contribute to protein folding: the hydrophobic effect and hydrogen bonding lead to protein stabilization, while the loss of conformational entropy and rotameric strain lead to destabilization. Another paper by Walshaw and coworkers explores the structurally interesting family of coiled-coil protein architectures, focusing on interactions

that lead to the formation of α -cylinders and α -sheets. The potential role of novel coiled-coiled structures in protein engineering is also discussed.

Finally, the book closes with two papers discussing the current state – and future potential – of applied protein design. The first of these by McPherson and Harrison, applies a variety of protein design techniques to the creation of a highly efficient protease inhibitor, with the goal of deploying this novel protein in plant roots to combat nematode pests. This paper compares the costs and potential benefits of rational design, in which a protein active site is modeled and then modified to change protein function, with directed evolution, in which a large number of protein variants are randomly generated, then tested in an effort to find functional modifications that serve the needs of investigators. This paper discusses the process – and eventual success – of rational design applied to the plant-nematode system, which used the rice seed cysteine protease inhibitor as a starting point to design a protease inhibitor with greatly increased potency. This process is then compared with directed evolution techniques. Potential techniques for generating diversity, including random mutation and DNA shuffling, are discussed in detail; the difficulties associated with testing and characterizing a large number of protein variants are also explored. The second paper dealing with protein design applies the process to the degradation of explosives by nitrate ester reductases. This study uses nitrate ester reductases derived from *Enterobacter cloacae* as a starting point in the search for enzymes capable of efficiently degrading highly recalcitrant explosive compounds such as 2,4,6-trinitrotoluene, better known as TNT. The paper discusses the family of enzymes believed to have the potential to breakdown such explosives, the relevant reactions and active site architectures, and the potential for improved efficiency or these reactions through rational protein design and site-directed mutagenesis.

All in all, this collection provides both a broad and deep look at the current state of the field of protein folding. We are told not only what is being discovered, but also why we should care, and what the future might hold. The papers are well selected and ordered, and complement each other well; the level of detail is just about right for the biologically-inclined neophyte. Most papers are presented with the tone of reviews rather than original studies, eliminating extensive discussion of materials and methods and focusing instead on the conceptual underpinnings of recent developments. That said, certain papers stray into levels of detail and specificity unlikely to interest the vast majority of scientists. While such details could prove interesting for those intimately involved in the field, the general tone of the book seems aimed more toward readers with a solid knowledge of science but minimal specific knowledge of protein folding, rendering such detail superfluous. In general, though, this collection provides an excellent overview of the current state and future potential of the study of protein folding, and is certainly worth the money for anyone interested in a wide-ranging introduction to this fascinating field.