Available online at www.sciencedirect.com
ScienceDirect
Current Opinion in Structural Biology

The quest to understand protein folding

Editorial overview

LS Itzhaki and PG Wolynes

Current Opinion in Structural Biology 2008, **18**:1–3

0959-440X/$ – see front matter

DOI 10.1016/j.sbi.2008.01.001

LS Itzhaki

MRC Cancer Cell Unit, Hutchison/MRC Research Centre, Hills Road, Cambridge CB2 0XZ, United Kingdom
e-mail: lsi10@cam.ac.uk

PG Wolynes

Department of Chemistry and Biochemistry and Center for Theoretical Biological Physics, University of California, San Diego La Jolla, CA 92093-0371, USA
e-mail: pwolynes@ucsd.edu

In humanity's quest to interpret the phenomenon of life using the fundamental laws, achieving an understanding of protein folding has long stood out as a key intellectual step. By folding, the one dimensionally encoded information of the genome is somehow sculpted into dynamic three-dimensional structures that can then function in the context of the cell.

To generations of workers in biology, chemistry and physics, finding the key to this code that transforms one into three and four dimensions, has seemed like the search for the Holy Grail. Rather than yielding to the first attacks made by the scientific community, the protein-folding problem has come to obsess many of us. Recently, however, powerful experimental techniques and computation tools along with new conceptual frameworks have enabled us to attain a fundamental understanding.

The articles in the present issue document recent progress in uncovering the secrets of protein folding. These have turned out to be simple in outline and yet rich in detail. The folding of the simplest proteins can be understood in terms of a funneled energy landscape guiding the assembly of three-dimensional structures through Brownian motion. Owing to the funneled landscape, the kinetics of folding the smaller proteins follow extraordinarily simple patterns in the laboratory, dominated by native stability. The elementary events of protein assembly are beginning to be worked out using highly detailed atomistic computer simulations in concert with experiment. At the same time, how these elementary events are orchestrated to give a final structure has been shown to be predictable using simulations with reduced models that average over many degrees of freedom, allowing the topology-determining features of the sequence to be laid bare. Controversial issues about the ensemble nature of folding pathways are on the verge of direct resolution in the laboratory using emerging single molecule techniques. Following on from this progress towards an understanding of folding *in vitro*, key aspects of folding in the cell are also being clarified. Such understanding also raises the prospect of intervention in the folding process as a therapeutic strategy.

Daggett and Fersht review the progress made by using all-atom models to simulate the more rapidly folding proteins. Thirty years ago accounts of folding emphasized the slowness of the process. Direct simulations over even a few milliseconds remain beyond today's readily available computational capabilities. Yet the search for the fastest folding proteins near the Eaton speed limit has uncovered simple small proteins that fold in a few microseconds, a time scale overlapping current direct all-atom simulations. Such computational studies, as well as inferences from rapid unfolding simulations at atomic detail, seem to agree reasonably well with experiment. Many of the early events are, however, sensitive to details of the potential model.

Clementi reviews recent progress in understanding the complete assembly of proteins by using reduced models that do not simulate the motion of every atom. These studies allow access to much longer time scales with statistically significant accuracy and good sampling. Achieving an understanding of folding from simple models would be hopeless were it not for the fact that the overall organization of the energy landscape of natural proteins is a funnel. The funnel-like organization that probably evolved to confer robustness to mutation also confers a robustness of the prediction of the folding mechanism to errors or simplifications in modeling. Clementi documents how a variety of features of folding can be inferred by simulating simplified models that are based on knowledge of the native structure with a little thoughtful consideration about the heterogeneity of the effective solvent-averaged interactions between residues.

Any process of organization by definition must involve ensembles of possibilities otherwise it is not an organizational act! Nevertheless the precise role of diversity in folding ensembles has remained one of the main sources of contention in the community of physical biochemists. Shuler and Eaton describe how increased experimental sensitivity has made possible the use of fluorescent probes to examine folding of protein molecules one at a time. In this way the ensemble averages can be picked apart and analyzed, allowing diversity of configuration to be quantified. Limited temporal and low structural resolution still prevent such methods from allowing the multiplicity of pathways to be completely examined but the ranges of early and late stage structure are coming into view.

The repetitive sequences of some proteins give them a quasi-one-dimensional structure in which the cooperative aspects of the folding process can be rendered rather malleable because of the paucity of three-dimensional contacts. Barrick, Ferreiro and Komives review the significant progress that has been made by confronting theory with experiment for one class of repeat proteins—the ankyrin repeats. By their contrast in behavior from more completely three-dimensional globular proteins, these repeat proteins provide unique opportunities for testing the underlying principles behind protein energy landscapes.

In vivo, doubtless many folding and unfolding events are going on at all times without the intervention of cellular machinery. Folding, however, is apparently too important for life, to have been left entirely unregulated within the cell. Many of the larger multi-domain proteins and multiprotein complexes, especially, are assembled and sometimes disassembled by subcellular machines. Often an orchestrated sequence of several different machines is required to guide a protein from its birth through the multiple stages of its folding to trafficking between various cellular locations, and finally to its destruction. Saibil reviews the great progress that has been made in understanding how one class of subcellular machines, the molecular chaperones, proofread and catalyze these processes. Structural studies that have caught proteins in the act of their chaperoned folding are giving real insights and are suggesting how chaperones might work to prevent misfolding and aggregation.

Assisted unfolding of proteins is often an early step towards their ultimate destruction. Protein degradation is not just housekeeping. Controlled destruction by the proteasome of short-lived regulatory proteins is a key part of the control of a myriad of cellular pathways, and is now understood to play as important and ubiquitous a role as protein phosphorylation. Errors in proteasome-mediated degradation are associated with many diseases including neurodegenerative disorders and cancer. Inobe and Matouschek review the role of folding and unfolding processes in targeting specific proteins for regulated degradation. Studies, including those using the more simple prokaryotic analogues of the proteasome, are beginning to reveal the multiple factors that can determine a protein's susceptibility to destruction.

The possible medical impact of understanding folding has recently come into view. Folding is so crucial to life that its wholesale failure would lead to complete nonviability of the organism and misfolding of even a few specific proteins can lead to disease. The most notorious of such misfolding diseases are caused by prions. Perrett and Jones review what we are finding out by studying prions in yeast. Yeast provides a more manageable system for genetic manipulation than do animals or man, allowing much to be learned. Prion propagation seems to be intimately connected with chaperone function. Structural and genetic studies are now beginning to help us to understand the atomic basis for the existence of strains, an otherwise perplexing example of protein-based, rather than nucleic acid-based, inheritance. A further important goal will be to resolve the structural changes that occur in the conversion of the native state of a protein into its prion form.

The diverse ensemble of structures that proteins adopt before they fold makes the usual approaches for finding drugs to intervene by binding, difficult to apply to misfolding diseases. Likewise interference in the folding of a pathogen's protein is a recent pharmacological strategy that presents new challenges for drug design. Broglia, Levy and Tiana show how understanding the folding mechanism of HIV protease, a dimer, provides a new strategy for treatment. Peptides that interfere with protease folding have been discovered. These may provide leads to a new pharmacological approach for AIDS therapy that complements currently used drugs in the therapeutic cocktail that target enzyme active sites. Owing to the ensemble nature of folding, these potential drugs may also be

robustly active even after the protease mutates, an event that otherwise can lead to drug resistances.

According to legend, drinking from the Holy Grail was supposed to impart everlasting life. The first sips from the cup of understanding protein folding do not yet realize this goal. Yet the involvement of folding in so much of the plan of life guarantees longevity for the study of folding as part of molecular and cellular biology.