Available online at www.sciencedirect.com

SciVerse ScienceDirect

Current Opinion in Structural Biology

Folding and binding

Editorial overview

Jayant Udgaonkar and Susan Marquesee

Current Opinion in Structural Biology 2013, **23**:1–3

For a complete overview see the Issue

0959-440X/$ – see front matter,

http://dx.doi.org/10.1016/j.sbi.2013.01.002

Jayant Udgaonkar

Tata Institute of Fundamental Research, National Centre for Biological Sciences, Bellary Road, Bangalore 560065, India
e-mail: jayant@ncbs.res.in

Jayant Udgaonkar is a Senior Professor at the National Centre for Biological Sciences, Tata Institute of Fundamental Research, Bangalore. His laboratory studies protein folding, unfolding and misfolding using a large variety of biophysical methods, including time-resolved FRET and hydrogen exchange coupled to mass spectrometry. The focus of the work in his laboratory is on the study of how protein folding, unfolding and aggregation reactions commence, and on the study of the temporal and structural heterogeneity characteristic of these reactions.

Susan Marqusee

Division of Biochemistry and Molecular Biology, Marqusee Lab, QB3 Institute, 592 Stanley Hall #3220, University of California, Berkeley, CA 94720-3220, USA
e-mail: marqusee@berkeley.edu

Susan Marqusee is Professor of Molecular and Cell Biology, Division of Biochemistry Biophysics and Structural Biology at the University of California, Berkeley. Her lab focuses on studies of protein folding and dynamics. Currently she is investigating the sequence determinants of the conformational landscape using experimental techniques ranging from the mechanical manipulation of single protein molecules to hydrogen exchange and ensemble–based kinetics.

Much of the progress that has been made in the past fifty years in the study of protein folding mechanisms, as well as of the energetics of folding and binding, has come from the use of techniques that study the properties of ensembles of protein molecules. While such methodologies have their limitations, especially in that many of the probes ensemble-average over all forms of the protein present, their application has yielded most of the knowledge we have today on folding pathways, in kinetic as well as in structural terms. In particular, the delineation of folding pathways defined by folding intermediates, and the characterization of folding intermediates using hydrogen exchange measurements as well as transition states using ϕ-value analysis, have all depended on such kinetic measurements. Ensemble measurements using different spectroscopic methods remain essential for establishing the basic mechanism of folding of any protein.

An important early insight into the mechanism of protein folding from ensemble-averaging measurements came from the identification and characterization of molten globule forms of proteins, initially in equilibrium studies. Then, the kinetic molten globule found as an early intermediate on the folding pathways of many proteins, was shown invariably to be equivalent structurally to the equilibrium molten globule. The highly dynamic nature of this wet molten globule (WMG) and its native-like topology was recognized more than twenty years ago. The review by Baldwin and Rose brings into focus the more recent realization that in its WMG form, the protein may have sufficient native-like topology, even in the absence of the tight side-chain packing characteristic of the native state, for it to be able to bind to a specific ligand and to possess specific function. WMGs are typically observed to form before the rate-limiting step in folding, but may populate even after the transition state. In contrast, the dry molten globule (DMG), identified so far mainly from unfolding studies with a few proteins, is likely to form only after the rate-limiting step in folding. Unlike the WMG, the DMG is characterized by an unsolvated core, but like the WMG, it has loose side-chain packing. As Baldwin and Rose point out, if the DMG turns out to be as ubiquitous a late intermediate, as is the WMG an early intermediate, established models of folding, particularly 'two-state' models, may need to be reevaluated.

The subsequent review by Bhattacharyya and Varadarajan further analyzes side-chain packing in WMGs and DMGs. This review brings out the important point that although much is being learnt about packing in the WMG, especially with respect to the native state, very little is yet known about the extent to which native-like packing is absent in the DMG. For example, it is not known whether the core of the protein has achieved tight packing in the DMG, and whether all regions of the structure in the DMG are dry, or only the

core region is dry. The review ends by pointing out that packing defects in native proteins may have functional roles, and by describing a new way in which native-state packing can be studied by mutagenesis.

Insightful application of statistical mechanics to the study of protein folding dates back to about fifty years ago, and theoretical and computational models of folding have contributed much toward how experimental scientists think about folding. Nevertheless, until recently it was not obvious how to connect, in a quantitative manner, computational studies to experimental ensemble-averaging kinetic studies. The latter are invariably carried out in the presence of chemical denaturant while the former invariably use ill-defined (to the experimental scientist) folding temperatures. The review by Thirumalai and co-workers highlights recent advances in computational methods that allow denaturant-dependent folding and unfolding kinetics to be predictable in some cases. The review further highlights computer simulations that have identified a DMG in which the core is accessible to urea but not to water.

The advent of submillisecond kinetic measurements in the study of protein folding, about fifteen years ago, had the same profound impact as it had on the study of enzyme function more than thirty-five years earlier. The review by Lapidus focuses on the heterogeneity revealed by such measurements in the earliest steps of folding as well as that present in the unfolded state itself. In particular, the review discusses recent results that suggest that the protein chain may sample different conformations in the unfolded state in refolding conditions much slower than originally expected. An enduring challenge is to obtain direct high-resolution structural data about the earliest events in protein folding.

Single molecule measurements first made an impact on biology about forty years ago, when recordings of current through single ion channels in lipid membranes became possible. Their advent in the field of protein folding about fifteen years ago came in the form of two types of measurements: single molecule force-spectroscopy measurements of tethered protein molecules, and single-molecule spectroscopy measurements of protein molecules diffusing through a confocal volume. In contrast to most ensemble-averaging kinetic studies, single molecule measurements can potentially identify rare transient folding/unfolding events, can allow on-pathway or off-pathway roles for intermediates to be assigned with relative ease, and can better allow multiple pathways to be delineated. Single molecule studies are complementary to ensemble-averaging studies, and are perhaps best done after ensemble-averaging studies have established the basic kinetic mechanism. A challenge for single molecule measurements is to provide structural data in addition to the kinetic data.

The review by Schuler and Hofmann describes how single molecule spectroscopy studies of folding have developed since their advent about ten years ago. In initial studies, rate constants were obtained from ensemble-averaging over many single folding and unfolding times of many different protein molecules, as had been done earlier for single channel current measurements. As the review describes, recent developments include the coupling of microfluidic mixing to single molecule detection, the analysis of transition path times, and the investigation of fast dynamics in unfolded proteins. The review by Žoldák and Rief describes recent applications of single molecule force spectroscopy to probe the dynamics of proteins under mechanical load. In such studies, force acts like a mechanical denaturant. The pulling direction and reaction coordinate can be precisely tuned and the measured extension and kinetics can be used to map the energy landscape of a single protein molecule. In the near future we are likely to see studies combining this approach with the fluorescent approaches detailed by Schuler and coworkers, which will lead to an unprecedented view of a protein's energy landscape.

Just like the single molecule experiments, molecular dynamic simulations have the potential to monitor the fluctuations of a single protein as it navigates its folding pathway. These simulations have been limited in their ability to sample proteins of the size and time scale probed by experimental studies. The review by Lane and co-workers notes that a combination of advances in hardware, software, and modeling has started to bridge this gap. They outline developments toward the next challenge – data analysis. How do we analyze these data and compare them to the experimental observables? They describe developments in the use of Markov State Models that highlights the limitations of the ϕ-value analysis used in traditional ensemble experiments of folding.

Ultimately, the amino acid sequence dictates the energy landscape and folding of a protein. It is well known that proteins with high sequence identity or homologous proteins encode the same three-dimensional structure. But what about the energy landscape and folding pathways? Are some features dictated by topology, while others are a result of specific amino acid interactions? Clarke and coworkers look for such trends by reviewing the folding of structurally and evolutionary related proteins. Their results highlight the limitations of using comparisons between protein folds to extract general principles for folding. While the folding of some proteins seems resilient to changes in sequence, others are not. These differences in malleability of the folding pathway bring up interesting questions about the relationship between evolution and topology.

The final two reviews focus on the relationship between conformational dynamics and molecular recognition and

binding. Traditionally, the dissection of molecular recognition and binding is biased by structural (enthalpic) interactions. Wand challenges this view in his review of recent developments to probe side chain fluctuations in the native state. By using the fast dynamics monitored by NMR relaxation as a probe for side-chain entropy, Wand outlines a method for exploring the relationship between side-chain conformational entropy and molecular recognition. Raleigh and coworkers discuss the role of molecular recognition in amyloid formation by reviewing the biophysics of islet amyloid polypeptide (IAPP) and the molecular basis for the cytotoxicity of IAPP aggregates *in vivo*.

All together, the ten reviews on Folding and Binding in this issue of *COSB* bring out the important point that although the protein folding field has matured much over the past fifty years, new ideas continue to emerge and old ideas continue to reemerge. For example, thinking about molten globules continues to evolve. So do ideas about the forces that maintain the integrity of protein structure. The path forward is clear. Experimental methods need to improve in both temporal resolution and structural resolution, and also at both the single molecule and ensemble levels. Computer simulations have to move beyond corroborating experimental results, to informing experimental scientists about what details of the folding process are unimportant and not worth worrying about. Theory has to go beyond just accommodating and rationalizing experimental observations, to predicting better folding mechanisms, kinetics, and binding in a quantitative manner. Only then will our thinking about protein folding, which is shaped by advances in experimental, theoretical and computational methods, become clearer. There remains a lot to be done.