Folding and binding: implementing the game plan

Editorial overview

Alan R Fersht and Valerie Daggett

The basic game plan for understanding the pathway of protein folding is to solve the structures of all ground and transition states at atomic resolution, and measure the rate constants for their interconversion. Denatured states and intermediates are accessible by NMR spectroscopy, transition states by Φ-value analysis, and very fast reactions by relaxation methods. The experimental data can be combined with atomistic simulation to describe the whole pathway at atomic resolution. Single-molecule studies can give information on the variety of different processes that are averaged out in bulk measurements on ensembles. This collection of ‘Opinions’ is focused primarily on recent developments in experimental studies of folding and how they interface with simulation.

Denatured and intrinsically disordered states of proteins are interesting with regard to both their folding to give globular proteins and as intermediates or starting points for misfolding. Mittag and Forman-Kay review the experimental determination of the structures of both denatured states and natively disordered proteins using primarily NMR and small-angle X-ray scattering (SAXS). The information about such states is usually sparse. Vendruscolo reviews how such minimal information can be combined with simulation to give high-resolution structural detail.

Detailed structures of transition states for folding have provided the main experimental clues about the folding energy landscape, mainly using Φ-analysis. The first studies on barnase and chymotrypsin inhibitor-2 (CI2) gave a static picture of the transition state, but this rapidly progressed to a more dynamic picture with the discovery of Hammond and anti-Hammond effects, corresponding to movements along or perpendicular to the reaction coordinate, respectively. Also, barnase could be switched from multistate to pseudo-first order (apparent two-state) kinetics. Lindberg and Oliveberg have expanded on conventional Φ-value analysis to give a moving picture of steps in a pathway and to examine the extent of pathway malleability. The variability of a pathway appears to depend on the number and position of nucleation foldons, minimal folding units. Lindberg and Oliveberg conclude that, if there is just one foldon, the pathway is relatively robust and the transition state plasticity small, as observed for the archetypical two-state folder CI2. Multiple overlapping foldons can lead to multiple pathways whose importance and order of events are sensitive to mutation, and to folding intermediates that might be at high energy or accumulate. Brockwell and Radford argue that all folding pathways proceed via intermediates, echoing our own views. But when the energy barriers are very small and the intermediates are highly transient and differ very little from neighboring states, there will always be debate on what is an intermediate. It is becoming increasingly well established that there is a continuum of mechanisms for the folding of proteins.
small domains, from nucleation-condensation to framework, depending on the stability of the individual folders or modules.

Crucial to understanding folding pathways is quantitative knowledge of the rate constants for the interconversion of the different states. Increasingly in the past decade, protein folding has been subject to measurements on, first, the submillisecond and, now, the submicrosecond timescale. Not only do these measurements illuminate the rates of fundamental processes in the early stages of folding, but also the discovery of ultrafast folding proteins means that experiment and all-atom simulation can be followed on the same timescale. Dyer has reviewed the main ultrafast folding experiments and their combination with simulation, which is giving reliable descriptions of entire pathways. This area will continue to grow apace.

The above reviewers mention the importance of fundamental studies to the problem of protein misfolding and aggregation. Kodali and Wetzel provide a particularly scholarly and in-depth review of amyloid fibril formation, concentrating on the polymorphism and the role of intermediates in assembly. Just as it is very difficult to assign unambiguously whether intermediates are on- or off-pathway in simple protein folding, it can be even more difficult to analyze what is an obligatory intermediate in the kinetics of fibril formation. Furthermore, change of conditions, addition of small molecules and mutation can affect the final structures of the fibrils.

Single-molecule studies are growing in importance. Fluorescence resonance energy transfer (FRET) studies are not reviewed here and are relatively small in number, but they give unique information about the variety of states present at equilibrium and their rates of interconversion. There has been considerably more work on single-molecule studies of the mechanical unfolding of proteins. ‘Pulling’ using atomic force microscopy is particularly relevant for proteins whose mechanical folding has biological importance, and is reviewed by Forman and Clarke.

Protein folding is primarily a self-recognition process, whereby atoms within a polypeptide chain organize themselves to form, in the main, the lowest energy ensemble. Individual proteins then might interact with each other, to form homo- or hetero-oligomeric complexes for function. The network of protein–protein interactions is at the heart of systems biology. The same techniques that can be applied to study intramolecular recognition in folding can be applied to the intermolecular process of molecular protein–protein recognition. The first studies used protein engineering to map out binding energy and specificity; these studies predated such experiments on protein stability. Schreiber and colleagues show how much further the area has progressed by analyzing the molecular architecture of binding sites and how they are organized.

The foundations of our current approaches to understanding the folding and binding of proteins were laid some 20 years ago, stemming from the combination of recombinant DNA methods with biophysical techniques. With increasing sophistication of experiment, different approaches are clicking rapidly into place and we now have a very good qualitative understanding of folding that could hardly have been envisaged two decades ago. The parallel developments in all-atom simulation are making the area of quantitative prediction of folding and binding much closer to our grasp.