Free energy barriers in protein folding and unfolding reactions

Santosh Kumar Jha and Jayant B. Udgaonkar*

National Centre for Biological Sciences, Tata Institute of Fundamental Research, Bangalore 560 065, India

Protein folding and unfolding reactions are slowed down by free energy barriers that arise when changes in enthalpy and entropy do not compensate for each other during the course of the reaction. The nature of these free energy barriers is poorly understood. The common assumption is that a single dominant barrier (k_B T), describable in terms of a single reaction coordinate, slows down the structural transition, which then becomes an all-or-none transition. This assumption has allowed the empirical application of transition state theory which has proven to be remarkably successful in describing protein folding reactions. Not surprisingly, much effort, both experimental and computational, has focused on determining the native and non-native interactions that determine the properties of the transition state, in order to determine which residues play crucial roles on the folding and unfolding pathways. The alternative hypothesis is that many small (< 3k_B T) barriers distributed on the energy landscape slow down the structural transition, which then becomes gradual and diffusive. Experimental, theoretical and computational evidence supporting this alternative hypothesis for describing the folding and unfolding of at least some proteins, has gradually been mounting.

Keywords: Cooperativity, energy landscape, kinetics, protein folding and unfolding, transition state.

Proteins are the functional entities in all living systems. They perform numerous functions, including catalysis of chemical reactions, transport of ions and molecules, coordination of motion, provision of mechanical support, generation and transmission of nerve impulses, and control of growth and differentiation. To be functional, a protein needs to fold into a specific tertiary structure. It has long been known that the functional structure of a protein depends on the sequence of residues of nascent polypeptide chains – the precursors of the folded proteins which in biology play the role of Maxwell’s demons. In a very real sense, it is at this level of organization that the secret of life (if there is one) is to be found. If we could not only determine these sequences but also pronounce the law by which they fold, then the secret of life would be found – the ultimate rationale discovered!

In 1969, Cyrus Levinthal pointed out that a polypeptide chain of 101 amino-acid residues, with each residue capable of having at least three accessible conformations, would have to sample 3^{100} = 5 \times 10^{30} conformations in its search for a single native conformation. If it took 10^{-13} s, the time taken for a chemical bond to rotate, to sample each conformation, then it would take 10^{27} years for an unfolded polypeptide chain to complete the search for its native conformation. Hence, a polypeptide chain would not be able to fold to its unique three-dimensional structure on the biological timescale of a few seconds, by a random search of the available conformational space. This implies that there must be defined pathways, each a particular sequence of structural events, available for a protein to fold. Understanding the temporal sequence of events that occur during folding has been a major challenge for experimental biochemists.

Inside a living cell, a protein exists in various conformational states. After synthesis on the ribosome it exists as an ensemble of unfolded conformations, and acquires a unique native structure by folding via various intermediate conformations. These conformational states of the protein exist in dynamic equilibrium with each other, and the population of each state depends upon the environmental conditions prevalent inside the cell. For example, under certain cellular conditions, large-scale structural fluctuations in the protein structure can lead to the formation of partially unfolded and misfolded forms, which have been shown to be the precursors for the formation of well-organized fibrillar aggregates. Hence, it is not only important to understand the forward reaction, i.e. how proteins fold, but also the reverse reaction, i.e. how they unfold, and to understand the nature of the free energy barriers that slow down protein folding and unfolding reactions.

Detailed characterization of the folding and unfolding pathways of proteins also has immense practical significance. It is expected that knowledge of the rules govern-

*For correspondence. (e-mail: jayant@ncbs.res.in)
ing protein folding and unfolding reactions will enable the design of proteins with desired stability and functionality. It will also help in engineering desired functionalities into existing proteins using recombinant DNA technology. Based on the knowledge gained from the structures and folding pathways of many proteins, attempts to produce new enzymes are already in progress\textsuperscript{10–14}.

One of the most poorly understood aspects of protein folding and unfolding reactions is the nature of the free energy barrier(s) separating the native (N) and unfolded (U) states. Many protein folding and unfolding reactions have been described as cooperative ‘two-state’ N \(\leftrightarrow\) U transitions\textsuperscript{15–17}, which implies that native structure forms or dissolves in a concerted, all-or-none manner analogous to a first-order phase transition\textsuperscript{18,19}. On the other hand, there is growing evidence which suggests that protein folding and unfolding transitions may be so highly non-cooperative, that they occur in many steps\textsuperscript{20,21} or even gradually\textsuperscript{22–28}. In this review, the current status of knowledge about the nature of the free energy barrier(s) which protein molecules traverse during folding or unfolding is presented. The degree of cooperativity accompanying the main structural transition, but not of the earliest events\textsuperscript{29}, is discussed. Experimental methodologies, which can measure folding and unfolding reactions at the single-residue level, and which have contributed immensely to our current knowledge, have also been discussed along with their applications. Kinetic studies of protein unfolding which reveal the nature of events following the rate-limiting step during folding have also been discussed briefly.

Models and theory of protein folding

\textit{Phenomenological models}

Experimental exploration of protein folding mechanisms is driven by the following conceptual models:

\textbf{Framework model:} This model envisages the folding reaction as the sequential formation of native-like micro-domains (\(\alpha\)-helices, \(\beta\)-hairpins, etc.). These native-like, small secondary structural units, are formed locally during the initial stages of protein folding, and come together by random diffusion and collision, which results in the formation of the final stable tertiary structure having native-like contacts\textsuperscript{30–37}.

\textbf{Nucleation and nucleation–condensation model:} According to the nucleation model, a few key residues of the polypeptide chain form a local nucleus of secondary structure in the rate-limiting step of folding. Around this nucleus, the whole native structure develops, as in a crystallization growth process\textsuperscript{38}. An extension of the nucleation mechanism is the nucleation–condensation model in which a nucleus of local secondary structure has poor stability by itself, and its stabilization requires interactions between non-local residues\textsuperscript{39}. Hence, the nucleation–condensation model envisages a diffuse folding nucleus, and all the secondary structure and native-like tertiary contacts form in a concerted manner in a single rate-limiting step\textsuperscript{39–42}.

\textbf{Hydrophobic collapse model:} This model posits that folding begins by an initial clustering of hydrophobic residue side chains which prefer to be excluded from an aqueous environment. The clustering of hydrophobic residues is expected to be non-specific and hence, to happen rapidly. The formation of an ensemble of collapsed structures, initially during folding, would drastically reduce the available conformational search-space\textsuperscript{41–49}. Hydrophobic residues of the protein are clustered in the interiors of the collapsed forms. The formation of secondary structure and consolidation of specific tertiary contacts is promoted in these collapsed conformations with relatively fluid structures\textsuperscript{50–52}.

\textit{Energy landscape theory}

Statistical mechanics-based models\textsuperscript{44,53–56} postulate that protein molecules traverse a funnel-shaped energy landscape during folding, and that protein folding pathways more closely resemble ‘funnels than tunnels in configuration space’\textsuperscript{57}. A folding funnel is a plot of the enthalpy against configurational entropy (Figure 1). An individual folding trajectory is envisaged for each polypeptide chain traversing down the folding funnel. Depending upon the...
asymmetry inherent in the energy landscape, large sets of the folding trajectories, with common features, may be averageable into folding pathways. Each such ‘macroscopic’ pathway would be distinguished by a specific progression of structural transitions, which is shared by the averaged trajectories. According to this viewpoint, intermediates are considered as kinetic traps which slow down the unfolding reaction, and transition from one ensemble of structures to the next on the unfolding pathway can happen on parallel routes.

Energy landscape theory largely ignores the diverse chemistry of the amino-acid residues building up a polypeptide chain. Consequently, the folding of a polypeptide chain appears to be opposed only by chain entropy. Energy landscape theory can accommodate many phenomenological observations made in protein folding studies, but it cannot necessarily predict them a priori. For example, it is difficult to predict protein folding rates, and how these rates, the contours of the free energy landscape, and indeed folding and unfolding mechanisms, may vary with changes either in the sequence of the polypeptide or in the folding conditions. An important utility of energy landscape theory may well be in delineating what details of polypeptide sequence and structure are not important in determining how a protein folds.

The different models of folding make different predictions about the nature of the free energy barrier(s) which protein molecules have to cross during folding. The nucleation and nucleation–condensation models predict that all the protein molecules pass through a unique transition state (TS) during folding. Computer simulations using off-lattice models suggest, however, that the critical nuclei envisaged by the nucleation and nucleation–condensation models, should be viewed as fluctuating mobile structures, thus implying non-unique transition states.

In contrast, the framework model predicts a hierarchical and progressive formation of protein structure, implying the existence of multiple transition states during folding. It also brings out the possibility that the ‘locking’ of secondary structural elements to form the final native fold (see above), could happen in different ways on a multitude of pathways, again implying that different protein molecules may cross different barriers during folding. The hydrophobic collapse model predicts that secondary structure forms after the collapse of the polypeptide chain. Non-native contacts may also be developed in the collapsed forms, as also predicted by energy landscape theory. In disagreement with the nucleation–condensation model, which considers the nucleus as an activated state, the initial collapse reaction of the polypeptide chain appears to be nearly barrier-less, as also predicted by energy landscape theory. In response to a change in external conditions such as temperature, pH or denaturant concentration is measured. An exponential or multi-exponential time-dependence of the change in a spectroscopic property is usually observed during folding and unfolding reactions. This

Nature of barrier(s) during protein folding

Protein folding reactions are usually described using the terminology and nomenclature that were established for small-molecule chemistry. They are, however, different from many other condensed phase chemical reactions in many significant ways. First, the structural transition of an unfolded polypeptide chain into a unique native fold involves the formation and breakage of many weak non-covalent bonds, in contrast to one strong covalent bond in classical chemical reactions. Secondly, protein–solvent interactions as well as solvent–solvent interactions (hydrogen bonds) play an important role during the folding of proteins. Thirdly, the size of the conformational ensemble changes dramatically. Non-polar amino-acid residues, which are solvent-exposed in the unfolded state, get ordered and buried in the hydrophobic core in the native protein. Water molecules, which had previously been ordered around non-polar side chains, become more mobile. Hence, the change in entropy associated with the change in ordering of water molecules plays an important role during folding, in addition to the change in configurational entropy. More recently, the importance of the contribution of the polar main-chain backbone (hydrogen bonds) in determining protein stability has been re-recognized. Thus, the thermodynamics of folding is defined by the delicate balance between the enthalpy and entropy of the protein–water system. The free energy barriers encountered by an ensemble of unfolded conformations as they proceed to the unique native state arise due to an incomplete compensation between the changes in entropy and enthalpy of the system, rather than due only to high-energy strained states. The dynamic nature of these barriers, and their thermodynamic and kinetic characterization, has remained a central focus of protein folding studies.

Kinetics of protein folding and diffusive nature of barrier crossing

It is necessary to study the kinetics of protein folding and unfolding reactions in order to determine the temporal sequence of events as well as the nature of the free energy barrier. Typically, the change in the reaction rate in response to a change in external conditions such as temperature, pH or denaturant concentration is measured. An exponential or multi-exponential time-dependence of the change in a spectroscopic property is usually observed during folding and unfolding reactions. This
The application of Kramer’s theory to introduce a significant amount of friction in protein dynamics to the motions of the solvent can residues in the interiors of proteins, and the coupling of multiple re-crossings of the barrier. In a diffusive process such as protein folding and unfolding, it is likely that the barrier is re-crossed multiple times before the reaction is complete. Moreover, the nature and meaning of the pre-exponential term is poorly defined when TST is applied to protein folding. Nevertheless, in the absence of easily applicable alternative models, TST continues to be used in the analysis of the results of kinetics studies of protein folding, despite having many shortcomings.

A more appropriate description of the folding reaction is given by a formalism introduced by Kramers, in which the role of Brownian motion or diffusive dynamics in barrier crossings is a key factor, and the possibility of multiple re-crossings of the barrier is taken into account. Kramers’ theory assumes that the diffusive motions of a protein molecule during folding are coupled to the motions of the solvent molecules, and this damping may significantly reduce the observed reaction rate compared to that predicted by TST. The diffusive nature of barrier crossing dynamics in protein folding reactions has been supported by computer simulations using lattice models. It has been shown that the dense packing of residues in the interiors of proteins, and the coupling of protein dynamics to the motions of the solvent can introduce a significant amount of friction in protein dynamics. The application of Kramers’ theory to folding reactions is, however, not straightforward, because it is difficult to determine experimentally how different dynamic modes of the protein are coupled to solvent fluctuations during folding. There have been some recent attempts, however, to measure experimentally the effects of friction on folding and unfolding dynamics by measuring the effect of external perturbations such as viscosogens, pressure and temperature on the kinetics of folding and unfolding. In many of these studies, the rate constants of folding or unfolding were seen to scale linearly with the inverse of the coefficient of viscosity of the solvent as predicted by Kramers’ theory, indicating a diffusive crossing of the barrier or unfolding barrier.

**Two-state folding**

**Structure of the TS and φ-value analysis**

The folding and unfolding reactions of many proteins have been characterized as cooperative ‘two-state’ transitions. This implies the existence of a unique TS (i.e. a single dominant free energy barrier) during folding and unfolding. The TS, by definition, is a hypothetical unstable state which lies at the top of the free energy barrier and hence, its structure cannot be characterized by direct experimental methods. Indirect methods based on linear free energy relationships used for determining the mechanisms of the chemical reactions of small organic molecules, like φ-value analysis, have been used routinely to study the structure of TS, and to map the fates of individual side chains in TS.

**Tertiary interactions and native-state topology**

In contrast to the relatively uniform distribution of the φ-values in TS for several small proteins, the distribution for many other proteins, including SH3 domains, barnase and CspB, has been found to be heterogeneous. For these proteins, regions of native-like interactions as well as relatively unstructured regions are present in TS, suggesting that the energetic perturbation of TS is proportional to that of the native state, for all of the residues investigated. This suggests that TS for these proteins resembles an expanded form of the native structure.
significant differences in the folding mechanisms (and presumably TS structures) were found for fatty acid-binding proteins, which are predominantly β-sheet proteins having the same fold and highly similar native structures\textsuperscript{112}. Mutational studies on many other proteins including Arc repressor\textsuperscript{110} and Rop\textsuperscript{113} have suggested that interactions which require specific alignment (for example, a buried salt bridge) may be difficult and energetically costly to achieve in the TS structure. Thus, it appears that tertiary interactions and topology may be important for assessing the determinants of folding rates.

**Relative contact order**

The importance of native-like topology in determining the folding mechanism has been shown by various computational and theoretical studies, in addition to a large body of experimental work as discussed above. The native-state topology of a protein is usually quantified using a parameter called the relative contact order (RCO), which is defined as the average sequence distance between all pair-wise contacts normalized by the number of residues\textsuperscript{114}. For many small, single-domain proteins which appear to fold in a ‘two-state’ manner, a strong correlation between folding rates and RCO has been found\textsuperscript{114}. Proteins with a lower RCO (such as helical proteins) fold faster compared to proteins possessing a high RCO (such as β-sheet proteins). This correlation implies that helical segments in a protein would fold faster than β-sheet regions. It is surprising then that within structurally similar proteins there exists a sizable variation in refolding rates\textsuperscript{17}. It appears that factors other than topology and tertiary interactions play a significant role in determining the TS structure during folding.

**Circular permutation**

The influence of RCO in TS of folding has been examined experimentally for many proteins, including T4 lysozyme\textsuperscript{115}, α-spectrin SH3 (ref. 116), RNase T1 (ref. 117), CI2 (ref. 118) and ribosomal protein S6 (refs 119 and 120), by studying the kinetics of folding and unfolding of circularly permuted forms of these proteins. In circular permutant forms of a protein, the order of secondary structural elements is re-arranged by joining the –N and –C termini using a peptide segment, and introducing new termini in different regions, so that a similar native fold and stability (similar enthalpic interactions) is retained in all the permutants. For many of these proteins, the folding nucleus is retained in the circularly permuted forms\textsuperscript{118,119}, and lowering RCO by means of circular permutation increases the rate of folding. Different circular permutants of ribosomal protein S6, having different values of RCO, were, however, observed to fold with similar rates\textsuperscript{120}. For α-spectrin SH3 (ref. 116), various circularly permutated forms of the protein were seen to fold via different folding pathways. The use of circular permutation in conjunction with φ-value analysis has also indicated that activation barriers during the folding and unfolding of proteins can be broad, flat and malleable and hence, would appear different in different folding or unfolding conditions\textsuperscript{121–123}.

**Limitations of φ-value analysis**

Although φ-value analysis has been quite useful in determining the structure of TS at the level of individual side chains, many interpretational ambiguities have been related to its usage. In φ-value analysis, it is usually assumed that the unfolded state of the protein is similar to a random coil and hence, does not get affected by the mutation. This may not be a valid assumption\textsuperscript{124–126}. Residual structures (both native and non-native like) are found to exist in the unfolded states of many proteins\textsuperscript{127–132}. It has been shown that such residual structures can be modulated by a change in solvent conditions and by mutagenesis, and that such modulations affect the stability of the unfolded protein\textsuperscript{127,133}.

While elegant in its simplicity, φ-value analysis has inherent limitations, implicit in relating thermodynamics directly to structure, and the method may be prone to experimental uncertainties\textsuperscript{108,134}. Furthermore, the meaning of partial φ-values, which have been commonly observed for most of the proteins studied using this method\textsuperscript{135}, also remains controversial\textsuperscript{134,136}. Although commonly interpreted as partial structural formation in TS, partial φ-values can also arise if TS is an ensemble of multiple structural forms, which are presumably formed on parallel pathways (it should be noted that in φ-value analysis, the interpretation of the data is based usually on the assumption of a single folding pathway). Hence, the structural interpretation of φ-value analysis remains ambiguous. In almost all cases where φ-value analysis has been reported, only one spectroscopic probe has been used to monitor the folding kinetics. In many studies, however, it has been seen that different probes show different folding kinetics (see below), indicating that the interpretations of φ-values determined using only a single probe may be unreliable.

**Multi-state folding**

**Theoretical studies**

A protein folding or unfolding reaction is governed by a free energy surface of high dimensionality and complexity because of the involvement of a large number of degrees of freedom\textsuperscript{9,137–139}. The multi-dimensional nature
of the potential energy surface describing a protein folding reaction, although not clearly implicit in the free energy reaction coordinate diagrams, is brought out fully by computational studies involving lattice models and energy landscape theory. \(^{19,160-162}\) Folding funnels and other multi-dimensional representations of potential energy versus conformation have highlighted the roughness and traps on the energy surface. They have also attempted to convey the interplay between the changes in entropy and enthalpy which occur during the course of folding.

**Experimental characterization**

Experimental characterization of the degree of cooperativity of the structural transition accompanying the folding reaction of a protein has been limited mainly due to the low sensitivity of the probes used to study them. Optical probes such as circular dichroism (CD), fluorescence, absorption spectroscopy, etc. are generally used to monitor folding and unfolding reactions, but they give information only about the average properties of all the conformational states of a protein present at the time of measurement, and do not reveal anything about specific structural changes happening in different parts of the protein. Thus, folding or unfolding reactions appear cooperative when measured using ensemble-averaging probes, and the heterogeneity of the system remains unresolved. In principle, the use of multiple probes in tandem, reporting on different structural changes which occur during the folding of a protein, can help in resolving the heterogeneity of the structural transition. Surprisingly, the folding kinetics of most ‘two-state’ proteins has been studied using only one or two probes, which can be misleading. \(^{145}\)

**Use of multiple probes in tandem**

In many cases where multiple probes are used to monitor the kinetics of the folding or unfolding transition, heterogeneity in the measured rates has been observed. The major folding phase of barstar has different rates even on the slow timescale. \(^{144,151,152}\) In some cases, the heterogeneity observed in the unfolding of many proteins \(^{161-165}\) has also been observed in a few cases. \(^{158,159}\) Non-exponential kinetics for protein folding reactions may also be the consequence of folding protein molecules confronting a distribution of free energies instead of a single free energy in the activation barrier to be surmounted. \(^{149}\)

**Use of residue-specific probes**

A full understanding of the degree of the cooperativity inherent in protein folding kinetics demands a complete description of the events happening at the individual residue level. Several experimental methods give direct residue-specific information about folding reactions: (a) real-time NMR methods, (b) pulsed hydrogen exchange methods (pulsed-HX) coupled with NMR, (c) pulsed cysteine-labelling methods (pulsed-SX), and (d) fluorescence resonance energy transfer (FRET) methods. Real-time NMR techniques, despite having the advantage of offering atomic-level resolution, suffer from low sensitivity and have been restricted to slow folding and unfolding reactions. Nevertheless, in some cases, they have indicated that different parts of a protein fold or unfold at different rates even on the slow timescale. \(^{144,151,152}\) Pulsed-HX experiments coupled with NMR detection allow measurements of folding on the ms timescale, but provide residue-specific information only on the main chain. The use of the pulsed-HX method to monitor the folding of several proteins \(^{12,71,146,147,153-156}\) has revealed that folding is heterogeneous and non-cooperative. Similarly, HX studies of the unfolding of many proteins \(^{157-160}\) have indicated that unfolding too is heterogeneous and non-cooperative. Surprisingly, this methodology has yet to be applied to the study of the folding of the apparent ‘two-state’ folders.

**Use of the pulsed-SX methodology**

In contrast to the pulsed-HX experiments, the pulsed-SX methodology provides direct structural information on the fate of individual side chains during folding, and has been shown to be an excellent probe for studying the change in structure during the folding and unfolding reactions of several proteins, at the level of individual side chains. \(^{161-165}\) In brief, side chains located in different parts of the protein structure are mutated to cysteine, one at a time, and the solvent accessibility of the individual cysteine thiol group to rapid chemical labelling is measured at different times of folding or unfolding. The extent to which a particular cysteine residue is involved in structure formation at any time of refolding is reflected by the
fraction of molecules in which the cysteine thiol gets labelled at that time.

The pulsed-SX method was used to show that the refolding of apomyoglobin starts with a collapse of the polypeptide chain in which side chains located in different parts of the protein are buried differentially\textsuperscript{162}. The main folding reaction also appeared to occur in a non-cooperative manner\textsuperscript{162}. In a separate study, this method also helped in the identification of the site for the initial tertiary structure breakdown during the unfolding of apomyoglobin\textsuperscript{166}. In these studies, however, the quantification of labelled and unlabelled protein in a sample required the cumbersome and problematic precipitation of the protein using trichloroacetic acid\textsuperscript{162,166}.

In an elegant extension of this methodology, the pulsed-SX experiment was coupled to mass spectrometry to determine the fractions of labelled and unlabelled proteins at different times of folding of barstar\textsuperscript{165} (Figure 2). The rates of burial of the cysteine thiols located at ten different locations in the protein were measured (Figure 2 a--d). A three-fold dispersion in the rates of cysteine thiol burial at different structural locations was seen during folding (Figure 2 e), which appeared to be equal to or three-fold faster than that measured by the change in fluorescence of the sole tryptophan residue present in the protein (Figure 2 e). The observation of a dispersion, albeit small, in the relative rates of burial of side chains located in different parts of the protein (Figure 2 f) is important as it suggests that the packing interactions necessary for the stability of the native protein develop in multiple steps during folding\textsuperscript{165}.

Equilibrium and pulse labelling of cysteine thiols have also been used for characterizing unfolding transitions under both low- and high-denaturant conditions for barstar\textsuperscript{163,164}. It was shown that native barstar can sample the fully unfolded conformation even in the absence of denaturant\textsuperscript{164}, and that competing pathways are available to the protein for unfolding\textsuperscript{163}.

---

**Figure 2.** The major (ms) refolding reaction of barstar was studied using the pulse-thiol labelling methodology in conjunction with mass-spectrometry. 

- **a.** Locations of side chains in the protein structure that were mutated to cysteine one at a time. 
- **b, c.** The sole thiol group in each protein was labelled with a short pulse of labelling reagent MMTS at different times of folding and the extent of labelling was quantified using mass-spectrometry. 
- **d.** Kinetics of the change in cysteine accessibility during the refolding of the Cys3 mutant form of barstar (having a single thiol residue at position 3 in the sequence) in 0.6 M urea at pH 9.2. Comparison of the fluorescence and cysteine accessibility-monitored apparent rate constants of fast refolding in 0.6 M urea at pH 9.2. 
- **e.** The observed fluorescence-monitored (empty bars) and cysteine accessibility-monitored (filled bars) refolding rate constants for the indicated mutant proteins. 
- **f.** The ratios of the cysteine accessibility-monitored rate constants to the corresponding fluorescence-monitored rate constants. Adapted, with permission, from Jha and Udgaonkar\textsuperscript{165}. 

---


Use of steady-state FRET

Although site-specific information is available from HX and SX measurements of folding or unfolding, they do not give much information about residues which are solvent-exposed in the native state. Unlike HX and SX experiments, FRET measurements can provide information about changes in specific intra-molecular distances involving both buried and exposed residues, and has proven to be a sensitive tool to monitor structural transitions in proteins. However, reliable structural mapping of the folding or unfolding pathway of a protein requires an extension of the FRET technique, from measurement at a single site to measurement at multiple sites in the system.

Both single-site and multi-site FRET measurements have been proven to be of great utility for characterizing the heterogeneity and cooperativity of protein folding and unfolding reactions. For example, an early application of steady-state FRET showed that the unfolding of yeast phosphoglycerate kinase occurs in multiple steps. An intermediate was shown to be populated on the folding pathway of engrailed homeodomain. By coupling FRET with ultra-rapid mixing methods it has been shown that a collapsed intermediate is formed early during the folding reaction of acyl-CoA binding protein. It is important to note that earlier equilibrium and stopped flow kinetic studies had indicated that the unfolding of acyl-CoA binding protein could be described by a ‘two-state’ mechanism. FRET measurements have been particularly useful in the study of the folding and unfolding of barstar, where they have shown that folding commences by an initial hydrophobic collapse, that the initial collapse reaction is a gradual structural transition, and that surface expansion occurs independently of core solvation during unfolding. Interestingly, it was also shown that the otherwise spectroscopically silent cis–trans proline isomerization reaction can be directly monitored by FRET measurements during unfolding. Although a great wealth of information is available from steady-state multi-site FRET measurements, they give an ensemble-averaged value of each individually measured distance, and cannot reveal much about the conformational heterogeneity in an ensemble.

Use of fluorescence anisotropy

Steady-state and time-resolved fluorescence anisotropy is another important method which can measure changes in molecular dimensions, during the folding or unfolding of a protein. The use of time-resolved fluorescence anisotropy decay measurements has shown that the consolidation of the hydrophobic core precedes substantial formation of specific structure during the refolding of barstar. This result is important as it implies that the rigidification of the core plays a major role in limiting the rate of the folding reaction.

Folding and unfolding through a continuum of intermediate forms

It was suggested many years ago, based upon statistical mechanical treatments of folding and unfolding reactions, that proteins might fold or unfold in a continuous manner. Energy landscape theory for protein folding predicts that intermediates are ensembles of structurally distinct forms, and describes TS ensemble as a collection of high-energy conformations. One of the major outstanding issues in protein folding concerns the experimental characterization of the structural heterogeneity of TS ensembles and the role of this heterogeneity in determining folding pathways.

The question really is whether there is effectively only a single dominant free energy barrier (of $\geq 5 k_B T$), describable in terms of a single reaction coordinate, present between the native and unfolded states, or whether there exists a distribution of small barriers (of $\sim 1–2 k_B T$) (Figure 3). In the first scenario it is expected that only two types of population distribution (native and unfolded-like molecules) will be present under any condition of folding (Figure 3 a). In the alternative scenario, since the different states are separated by small energy barriers, a continuum of intermediate forms is expected to be populated, and the population distribution of different intermediate forms is expected to change continuously with a change in folding conditions (Figure 3 b). Experimentally distinguishing between these two possibilities remains a challenge because most of the techniques used to measure the folding or unfolding reaction give ensemble-averaged values of the physical quantities measured, and do not give any information about the distribution of the physical quantity over different members of the ensemble. For example, it has been difficult to establish unequivocally whether the small protein BBL is a ‘two-state’ folder or whether it folds in a downhill manner through a continuum of intermediate forms.

Use of single-molecule and time-resolved FRET

Recently, the use of high-resolution probes like time-resolved fluorescence resonance energy transfer (TR-FRET) and single-molecule fluorescence resonance energy transfer (sm-FRET) methodologies, which can distinguish between different structural forms present during a folding or unfolding reaction on the basis of the difference in the distributions of intra-molecular distances, has revealed the highly non-cooperative nature of protein folding and unfolding reactions. The use of sm-FRET has shown that RNase H unfolds in a gradual
Figure 3. Energy diagrams for ‘two-state’ (a) and continuous (b) protein unfolding scenarios. In the ‘two-state’ unfolding scenario, one dominant free energy barrier between the N and U states ensures that only these two forms are populated either under different conditions or at different times of unfolding. In the alternate scenario, the unfolding reaction is mediated by a large number of small distributed barriers (~1–2 $k_BT$). This leads to gradual changes in the structure of the protein, and the single population of molecules changes gradually with a change in unfolding conditions or at different times of unfolding.

According to TST, the kinetics of an elementary step during a chemical reaction is defined by the waiting time (attempt frequency), whereas the actual transition time (barrier-crossing time) over the barrier is too fast to observe. Sm-FRET studies appear ideal to test this basic tenet of TST. Recent sm-FRET studies have put an upper time limit of ~200 $\mu$s for the barrier crossing time, for a folding or unfolding reaction. Some studies have, however, also brought out the possibility that the transition between the two energy states during the folding or unfolding reaction of a biomolecule may not occur in a ‘sudden jump’ fashion, but might occur in a gradual manner over many seconds. Single-molecule fluorescence studies of the slow unfolding reaction of green fluorescent protein have shown that each protein molecule jumps continuously between many conformational sub-states for many milliseconds, immediately before flipping to the U state. Hence, these studies seem to indicate a folding scenario with no defined kinetic barrier between the unfolded and folded states. Sm-FRET measurements suffer, however, from low time resolution (the fluorescence is averaged over millisecond bursts, and the distribution is obtained by looking at many different molecules), and in general it is not possible to observe the same molecule over a prolonged time duration due to technical reasons.

In contrast to sm-FRET, TR-FRET-based estimation of the distance between the fluorescence donor and acceptor is done by measuring the fluorescence lifetime of the donor in the absence as well as in the presence of an acceptor, and has much better time resolution. The extent of quenching of the fluorescence lifetime of the donor in the presence of the acceptor is determined, and is related to the distance between the donor and acceptor by Forster’s relation. In a biomolecular system, the distribution of the distances between the donor and the acceptor results in a distribution of energy transfer rates which can be measured as a complex fluorescence intensity decay of the donor. Generally fluorescence lifetimes are of the order of a few nanoseconds. The structural transitions between the native and unfolded states are slower than the donor fluorescence lifetime, and hence, TR-FRET-based measurements offer ‘snap-shots’ of population distribution rather than a weighted-average. Thus, such measurements yield the distribution of donor lifetimes, and subsequently the distribution of donor–acceptor distances, corresponding to the conformations sampled in the system. TR-FRET measurements have been informative in determining conformational heterogeneity during protein folding or unfolding, and in unfolded proteins. Denaturant-dependent, non-random structure was shown to be present in the unfolded state of barstar. Interestingly, in a separate study, it was shown that under conditions where ensemble-averaged probes suggested ‘two-state’ unfolding of barstar (Figure 4a), TR-FRET measurements indicated that the structure...
Figure 4. Structure is lost in a progressive manner during the unfolding of (a–c) barstar and (d–f) BBL. a, Change in the fraction of unfolded protein with different concentrations of urea as calculated from measurements of the fluorescence intensity at 360 nm (open circle) and the ellipticity at 222 nm (open square) is similar. The distribution of intra-molecular distance between Trp53 and a thionitrobenzoate (TNB) moiety placed at Cys82 of barstar, however, changes continuously with the concentration of urea. The fluorescence lifetime of tryptophan increases with increase in the distance from the TNB moiety. b, c, Fluorescence lifetime distributions of Trp53, in a single tryptophan (Trp53) and single cysteine (Cys82) containing form of barstar in which the sole thiol is labelled with TNB. b, 0 M urea (solid line), 1.8 M urea (dashed line), 3.2 M urea (dotted line) and 3.6 M urea (dashed–dotted line). c, 3.7 M urea (solid line), 4.1 M urea (dashed line), 6 M urea (dotted line) and 8 M urea. a–c. Reproduced with permission from Lakshmikanth et al. 25. d–f, Thermal unfolding of BBL measured atom-by-atom using NMR. d, Plot of chemical shift against temperature for nine representative protons. e, Histogram of the values of the denaturation mid-point temperature, $T_m$, for all 158 protons monitored. Protons displaying three-state behaviour (for example, green curves in d) provide two $T_m$ values to the histogram. $<T_m> = 303.7$ K; $\sigma_{T_m} = 16.9$. f, Comparison of the low resolution (circular dichroism, red) thermal unfolding curve with the average of the 158 normalized atomic unfolding curves (NMR, blue). The $y$-axis represents the amplitude of the second singular value for the circular dichroism spectra versus $T$ (red), or for the matrix of 158 NMR chemical shifts versus $T$ (blue). (Inset) Derivatives of the curves. d–f. Adapted by permission from Macmillan Publishers Ltd [Nature], (Sadqi et al.) 28, copyright (2006).

is lost incrementally and not in an all-or-none manner (Figure 4b and c). TR-FRET measurements have also shown that conformational heterogeneity exists during the initial stages of the folding of cytochrome c, and TIM barrel protein. 219. TR-FRET measurements have also proven to be useful in showing that intermediates during folding or unfolding are ensembles of structurally distinct forms, and that different folding pathways dominate under different folding conditions, as predicted by energy landscape theory. 65. A late intermediate, which accumulates during the folding of barstar, was shown to be an ensemble comprised of different structural forms, some unstructured and others highly structured. 198. It was observed that a change in the conditions of folding, from more stabilizing to less stabilizing, not only reduced the extent to which the intermediate ensemble was populated, but also affected the structural composition of the intermediate ensemble. Under greater stabilizing conditions, the more structured members of the intermediate ensemble were preferentially populated; under less stabilizing conditions, the less structured forms were preferentially populated. The observation that the structure apparent in a folding intermediate depends on the conditions employed to study folding is important because it implies that the folding pathway observed for a given protein will appear different under different conditions and different free energy barriers will be crossed in different folding or unfolding conditions.

Use of other high resolution probes

NMR methods have also been useful in revealing the continuous nature of folding and unfolding reactions. The thermal unfolding of a GCN4-like leucine zipper was
shown to occur in multiple steps\textsuperscript{23}. NMR studies have indicated that gradual disruption of side-chain packing occurs during the pH-induced equilibrium unfolding of CHABII (ref. 24). In a recent experiment, the thermal unfolding of BBL was monitored using NMR\textsuperscript{28}. The chemical shifts of various protons located in different parts of the protein appeared to change in an asynchronous manner during unfolding (Figure 4\textit{d}). Also, there existed a wide distribution in the midpoints of the unfolding transitions monitored by different protons (Figure 4\textit{e}), indicating that the unfolding of this protein occurs in a highly non-cooperative manner. Interestingly, the ensemble-averaged change in chemical shifts matched the unfolding transition monitored by CD (Figure 4\textit{f}). Recently, the equilibrium unfolding of barstar was also studied using \textsuperscript{19}F-NMR, and the data suggested that the protein unfolds in many steps\textsuperscript{202} as had been indicated in a previous study\textsuperscript{25}.

The application of UV-resonance Raman spectroscopy and single-molecule force spectroscopy to study protein folding and unfolding reactions has also revealed that proteins indeed traverse rugged energy landscapes during folding. For example, UV-resonance Raman spectroscopy studies on Trp-cage, a small synthetic protein, indicate that equilibrium thermal unfolding is gradual and spatially decoupled\textsuperscript{26}. It is interesting to note that previous kinetic studies using low-resolution optical probes had suggested that this protein was an ultra-fast ‘two-state’ folder\textsuperscript{203}. The use of single-molecule force spectroscopy has revealed that ubiquitin acquires structure continuously and slowly (on the seconds timescale), in a gradual manner in multiple discrete stages during its folding\textsuperscript{204}.

**Importance of a rugged free energy landscape**

It is not surprising to note that several proteins appear to fold or unfold via a continuum of intermediate forms, by traversing rugged energy landscapes. The structures of native proteins are believed to be stabilized, to a large extent, by the sequestration of hydrophobic residues away from water in the protein core\textsuperscript{146}. Because of the non-specific nature of the hydrophobic interactions, alternative hydrophobic core-packing arrangements could exist in proteins\textsuperscript{205}. These alternative arrangements could be less stable than the packing arrangement in the native state, but more stable than that in the unfolded state. This would produce a rugged and dynamic folding free energy landscape with shallow energy wells and transition barriers, where the protein can explore many conformational states of similar energies but distinct structures. The dynamic ability of native proteins to switch between different conformations reversibly might be important for many functions such as transmission of signals within and between cells\textsuperscript{206}, changing interaction partners\textsuperscript{207} and ligand binding\textsuperscript{208}. For example, the dynamic structure of myoglobin is important for the oxygen molecule to reach its binding site, and hence, for the protein to perform its function\textsuperscript{208}.

The ability of proteins to fluctuate continuously between semi-stable states might be important for the acquisition of new traits during evolution\textsuperscript{209–211}. It has been shown recently that a computationally designed protein Top7, devoid of an evolutionary history, folds in a highly non-cooperative manner via a rugged energy landscape\textsuperscript{209}. This result indicates that cooperative folding via a smooth energy landscape, which has been observed for many small naturally occurring proteins, could be a product of natural selection. Here, it is important to note that computer simulations suggest that a protein may unfold either in an all-or-none fashion or in a gradual fashion, and can switch between the two mechanisms upon a small change in solvent conditions or the primary sequence of the protein\textsuperscript{212}.

The observation that folding or unfolding may occur via a continuum of intermediate conformations, also has important implications for understanding protein aggregation reactions that lead to the formation of amyloid fibrils associated with many diseases. This is because the formation of amyloid fibrils can proceed from different conformations of partially unfolded proteins\textsuperscript{213}.

**Nature of TS during unfolding**

**Importance of understanding protein unfolding**

For a complete understanding of the free energy landscape traversed by a protein during folding, it is important to understand the nature of free energy barriers and to obtain detailed structural information on folding intermediates, encountered by the protein both before and after the rate-limiting step (the highest barrier encountered by the protein during folding). Although protein refolding studies provide a wealth of information regarding the nature of the free energy barrier during folding, they provide limited information about the free energy barriers which are crossed after the rate-limiting step of folding. This is because the events following the rate-limiting step occur on the downhill side of the major barrier, and are too fast to be captured using traditional methods. Protein unfolding studies become a method of choice to get this information as the initial events during unfolding can be expected to be similar to the events that follow the rate-limiting step of refolding\textsuperscript{157}.

**Unfolding of proteins also occurs in multiple steps**

It is generally believed that intermediate structures do not populate during unfolding. When multiple structural probes were used, however, to follow the unfolding reac-
tions of several proteins, unfolding intermediates could be shown to be populated transiently in conditions that favour either the unfolded state or the native state. It has also been shown that even though a protein may unfold through multiple pathways, the transition states on the different pathways may be similar in energy even while differing significantly in structure. An immunoglobulin domain of titin was shown to switch between two unfolding pathways by a change in the unfolding conditions. It was shown that the highly compact TS of one unfolding pathway gets destabilized with an increase in the concentration of the denaturant, and that the major population of protein molecules shifts toward another pathway with a less structured TS. Recent unfolding experiments on the SH3 domain of PI3 kinase demonstrated the presence of an intermediate which was shown to be populated after the rate-limiting step of folding. It is important to note that an earlier study of the refolding of this protein could not detect the presence of any intermediate form and concluded that the protein folds in a cooperative ‘two-state’ manner.

Dry molten globule hypothesis

Two hypotheses have been widely discussed to describe the poorly understood nature of the rate-limiting step during the unfolding reaction of a protein. In the first hypothesis, the rate-limiting step during unfolding is the penetration of water molecules inside the hydrophobic core which results in large-scale conformational rearrangement of the protein backbone. The protein becomes unstable in denaturing conditions and overall unfolding occurs rapidly. The second hypothesis is the dry molten globule hypothesis, which asserts that unfolding begins with an expansion of the native protein under the influence of thermal forces (Figure 5). At a critical degree of expansion, the side chains become free to rotate. The disruption of the tight packing of side chains in the protein core leads to a dense TS, which does not allow the penetration of solvent molecule inside the hydrophobic core. This was postulated to be the rate-limiting step during unfolding. In the second step of unfolding, the dry molten globule becomes wet and swells gradually to become a random coil (Figure 5). There was, until recently, however, little experimental evidence supporting the formation of the ‘dry molten globule’ state initially during unfolding.

Experimental demonstration

The first clue that a dry molten globule might be populated during protein unfolding came from NMR
measurements of the unfolding of ribonuclease A. It was observed that a few side chains became free to rotate early during unfolding whereas the secondary structure of the protein remained intact as measured by far-UV CD. It was also shown in a separate study that protons residing in the core of native ribonuclease A were resistant to exchange with solvent protons in this intermediate, indicating that the core of the intermediate is dry. Similar results were obtained in studies of the unfolding of 6-19F-tryptophan-labelled dihydrofolate reductase. A non-native intermediate was seen to form early during unfolding, in which the tryptophan moiety is not hydrated and secondary structure of the protein is intact, but tertiary interactions are broken. 17O relaxation dispersion NMR experiments have also helped in the identification of equilibrium dry molten globules for three proteins. Measurements using optical methods have shown that an on-pathway equilibrium dry molten globule is populated during the salt-induced folding of the high-pH-unfolded form of barstar.

It was expected that the perturbation of close packing interactions in the dry molten globule would lead to significant movement of secondary structural elements away from each other. The detection of the rotation or translation of an α-helix or the fraying movement of a β-strand during the formation of a dry molten globule, which would constitute the most direct evidence in support of the dry molten globule hypothesis, had been difficult to capture in experiments, partly because of the limited usage of residue-specific probes to study protein unfolding reactions. There had also been virtually no experimental evidence validating the second tenet of the dry molten globule hypothesis that the swelling of the dry molten globule gets broken only after the entry of water molecules inside the hydrophobic core. Hence, the observation that a dry molten globule state is populated early during unfolding is important because it indicates that dispersion forces also play a major role in maintaining the integrity of the native structure. Furthermore, it also suggests that TS of unfolding is an expanded form of the native protein, as inferred in an earlier study of the unfolding of lysozyme.

Concluding remarks

It is commonly believed that hydrophobic interactions are of paramount importance in determining the stability of a protein fold. A dry molten globule is a relatively stable structure with perturbed close packing interactions but an intact secondary structure. The secondary structure of the dry molten globule gets broken only after the entry of water molecules inside the hydrophobic core. Hence, the observation that a dry molten globule state is populated early during unfolding is important because it indicates that dispersion forces also play a major role in maintaining the integrity of the native structure. Furthermore, it also suggests that TS of unfolding is an expanded form of the native protein, as inferred in an earlier study of the unfolding of lysozyme.
icates that at different times of folding or unfolding, the intermediate ensemble could be dominated by different sets of conformation with varying degrees of structure. In future, it is expected that the increased time-resolution of single-molecule methodologies, coupled to theoretical and computational studies, will shed more light on the nature of the free energy barrier(s) which proteins explore during their folding and unfolding reactions.


Pace, C. N., Polar group burial contributes more to protein stability than nonpolar group burial. Biochemistry, 2001, 40, 310–313.


REVIEW ARTICLE


228. by pulsed hydrogen exchange.
226. breakage in the process of unfolding of ribonuclease A measured
222. transitions in protein molecules. I. Why denaturation of globular
221. Tanford, C., Protein denaturation. C. Theoretical models for
222. Kiefhaber, T., Labhardt, A. M. and Baldwin, R. L., Direct NMR
221. Galani, D., Fersht, A. K. and Perrett, S., Folding of the yeast
220. Jamin, M., Yeh, S. R., Rousseau, D. L. and Baldwin, R. L., Sub-
219. Jha, S. K. and Udgaonkar, J. B., Direct evidence for a dry molten
globule intermediate during the unfolding of a small protein.
218. Wani, A. H. and Udgaonkar, J. B., Revealing a concealed inter-
mediate that forms after the rate-limiting step of refolding of the
217. Guijarro, J. I., Morton, C. J., Plaxco, K. W., Campbell, I. D. and
Dobson, C. M., Folding kinetics of the SH3 domain of PI3 kinase
by real-time NMR combined with optical spectroscopy. J. Mol.
216. Kiefhaber, T., and Baldwin, R. L., Kinetics of hydrogen bond
breakage in the process of unfolding of ribonuclease A measured
by pulsed hydrogen exchange. Proc. Natl. Acad. Sci. USA, 1995,
92, 2657–2661.
transitions in protein molecules. I. Why denaturation of globular
protein is a first-order phase transition. Biopolymers, 1989, 28,
1667–1680.
214. Zaidi, F. N., Nath, U. and Udgaonkar, J. B., Multiple intermedi-
213. Jamin, M., Yeh, S. R., Rousseau, D. L. and Baldwin, R. L., Sub-
millisecond unfolding kinetics of apomyoglobin and its pH 4
212. Otzen, D. E. and Oliveberg, M., Burst-phase expansion of native
protein prior to global unfolding in SDS. J. Mol. Biol., 2002,
315, 1231–1240.
211. Galani, D., Ferst, A. K. and Perrett, S., Folding of the yeast
prion protein Ure2: kinetic evidence for folding and unfolding
studies of the unfolding of thioredoxin indicate stabilization of a
partially unfolded, aggregation-competent intermediate at low
209. Patra, A. K. and Udgaonkar, J. B., Characterization of the fold-
ing and unfolding reactions of single-chain monellin: evidence
for multiple intermediates and competing pathways. Biochemis-
try, 2007, 46, 11727–11743.
208. Wright, C. F., Lindorff-Larsen, K., Randles, L. G. and Clarke, J.,
Parallel protein-unfolding pathways revealed and mapped.
207. Wani, A. H. and Udgaonkar, J. B., Revealing a concealed inter-
mediate that forms after the rate-limiting step of refolding of the
206. Privalov, P. L. and Makhatadze, G. I., Contribution of hydration
actions on the transition state of lysozyme unfolding. I. Effect of protein–solvent inter-
actions by stronger dispersion interactions with proteins than
water implies a 2-stage unfolding. Proc. Natl. Acad. Sci. USA,
205. Hoeldtli, S. D. and Frieden, C., Stopped-flow NMR spectro-
copy: real-time unfolding studies of 6-19F-tryptophan-labeled
USA, 1995, 92, 9318–9322.
204. Denisov, V. P., Jonsson, B. H. and Halle, B., Hydration of dena-
253–260.
203. Jha, S. K. and Udgaonkar, J. B., Direct evidence for a dry molten
globule intermediate during the unfolding of a small protein.
202. Jha, S. K., Dhar, D., Krishnamoorthy, G. and Udgaonkar, J. B.,
11118.
201. Hagen, S. J., Exponential decay kinetics in ‘downhill’ protein
200. Privalov, P. L. and Gill, S. J., Stability of protein structure and
234.
199. Privalov, P. L. and Makhatadze, G. I., Contribution of hydration
to protein folding thermodynamics. II. The entropy and Gibbs
198. Segawa, S. and Sugihara, M., Characterization of the transition
state of lysozyme unfolding. I. Effect of protein–solvent inter-
actions on the transition state. Biopolymers, 1984, 23, 2473–
2488.

ACKNOWLEDGEMENTS. We thank the members of our laboratory
for discussions. J.B.U. is a recipient of the JC Bose National Fellow-
ship from the Government of India. S.K.J. acknowledges the Council of
Scientific and Industrial Research, New Delhi, for a Shyama Prasad
Mukherjee Fellowship. Work in our laboratory has been funded by the
Tata Institute of Fundamental Research, the Department of Biotec-
nology, and the Department of Science and Technology, Government of
India.

Received 18 February 2010; revised accepted 1 July 2010