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 ($> 3\,k_BT$), 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 ($< 3\,k_BT$) 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 is coded for by its primary amino-acid sequence[1–5], but the mechanism by which this happens is still not completely understood. Understanding the mechanism of how proteins fold has great relevance to modern biology. Loosely transliterating the words of Jacques Monod[6,7]:

'The ultimate rationale behind all purposeful structures and behavior of living beings is embodied in 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 $3^{100} = 5 \times 10^{47}$ 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[4,8]. 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[4]. 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[9]. 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[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 ⇌ U transitions[15–17], which implies that native structure forms or dissolves in a concerted, all-or-none manner analogous to a first-order phase transition[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[20,21] or even gradually[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[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

Phenomenological models

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

Framework model: This model envisages the folding reaction as the sequential formation of native-like micro-domains (α-helices, β-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[30–37].

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[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[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[39–42].

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[43–49]. 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[50–52].

Energy landscape theory

Statistical mechanics-based models[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'[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
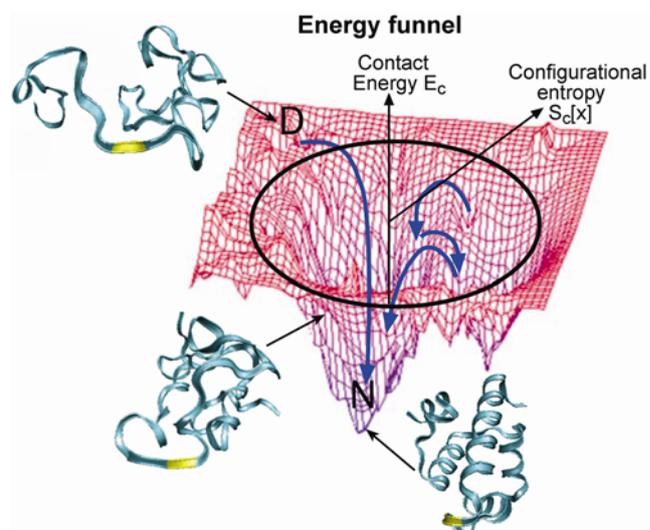

Figure 1. Funnelled folding. The unfolded conformation at high energy (D), compact globule at moderate energy, and the native state (N) at low energy are shown. The radial coordinate is proportional to the logarithm of the number of protein conformations at a given energy (the configurational entropy, which is higher at higher energy). The angular coordinate symbolizes the many other folding coordinates. Reproduced with permission from Gruebele[61].

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 folding reaction, and transition from one ensemble of structures to the next on the folding 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[58].

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 'docking' 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[33]. 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[59], 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[52,60–63].

Energy landscape theory suggests that the folding of proteins generally does not occur by an obligate series of clearly defined intermediates, defining a 'pathway'[64], but by a multiplicity of routes down a folding funnel. Hence, the folding process is described in terms of a progressive organization of ensembles of partially folded structures. Energy landscape theory also accommodates downhill folding, wherein proteins can fold without encountering any free energy barrier under certain thermodynamic conditions[54,61,65].

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[66]. 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[46,67], 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[68,69]. 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

observation has been interpreted usually to represent a single barrier or a few barriers along the reaction coordinate(s), rather than a distribution of barriers, and the kinetics data are analysed using transition state theory (TST). It should be noted, however, that TST, which is commonly used to describe the folding or unfolding kinetics of proteins[15,39,70,71], has only an empirical basis[72,73], and does not take into account the phenomenon 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[74], 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[75,76] 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. Kramer's 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[77,78]. 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[79–81]. The application of Kramer's 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[82–84]. 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 viscogens, pressure and temperature on the kinetics of folding and unfolding[85–94]. In many of these studies, the rate constants of folding or unfolding were seen to scale linearly with the inverse of the co-efficient of viscosity of the solvent[85,86,94] as predicted by Kramer's theory, indicating a diffusive crossing of the folding 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[17]. 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[95–97], 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[70,98–103]. In φ-value analysis, energetic interactions of a suitably mutated side chain in TS are compared to the energetic interactions of that side chain in the native state, relative to the unfolded state in both cases ($\phi = \Delta\Delta G_{TS–U}/\Delta\Delta G_{N–U}$). A φ-value of unity suggests that the side chain of the residue is in a native-like environment in TS, whereas a value of zero implies an unfolded-like environment.

The effects of point mutations on refolding and unfolding kinetics have been studied on many proteins, including barnase[70], CI2 (ref. 39), P22 *Arc* repressor[104], and src[105] and α-spectrin[106] SH3 domains. For many of these proteins, a linear relationship between $\Delta\Delta G_{TS–U}$ and $\Delta\Delta G_{N-U}$ has been observed, 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.

Tertiary interactions and native-state topology

In contrast to the relatively uniform distribution of the φ-values in TS for several small proteins[39,104], the distribution for many other proteins, including SH3 domains[105,106], barnase[107] and CspB[108] 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 TS is 'structurally polarized'[105,108]. The heterogeneity in φ-values exhibited by these proteins has been attributed to differences in topology between helical and β-sheet proteins. This is not surprising as helical proteins have been suggested to have a more 'delocalized nucleus' than β-sheet proteins[109]. It is interesting to note, however, that for many helical proteins, including *Arc* repressor[110] and monomeric λ repressor[111], point mutations can have significant effects on the folding mechanism, and can change the position of TS along the reaction coordinate.

The src and α-spectrin SH3 domains have similar native structures and φ-value distributions in TS, despite little similarity in sequence. It has been suggested, based upon this observation, that the topology of the protein rather than the specific amino acid content is the main determinant of the TS structure[105]. On the other hand, sig-

nificant 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[112]. Mutational studies on many other proteins including *Arc* repressor[110] and Rop[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[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[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[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[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 permutated forms[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[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[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[124–126]. Residual structures (both native and non-native like) are found to exist in the unfolded states of many proteins[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[125,133].

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[108,134]. Furthermore, the meaning of partial ϕ-values, which have been commonly observed for most of the proteins studied using this method[135], also remains controversial[134,136]. Although commonly interpreted as partial structure 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[19,137–139]. 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[19,140–142]. 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[143].

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 when measured by different probes, indicating the presence of multiple barriers and intermediate structures on the folding pathway of this protein[51,144,145]. Similar results have been observed for several other proteins, including lysozyme[146], cytochrome c[147], barnase[71] and ribonuclease A[32]. In some cases, the heterogeneity observed in the folding kinetics when measured by multiple probes has been interpreted as arising from the presence of many small barriers (and not one large barrier) separating the unfolded and native states, and as a signature of downhill folding[148]. Highly non-exponential kinetics (kinetics

following a power law dependence), similar to that observed in studies of the ligand binding and release reactions of myoglobin[149], has also been observed in a few cases[148,150]. 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[32,71,146,147,153–156] 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
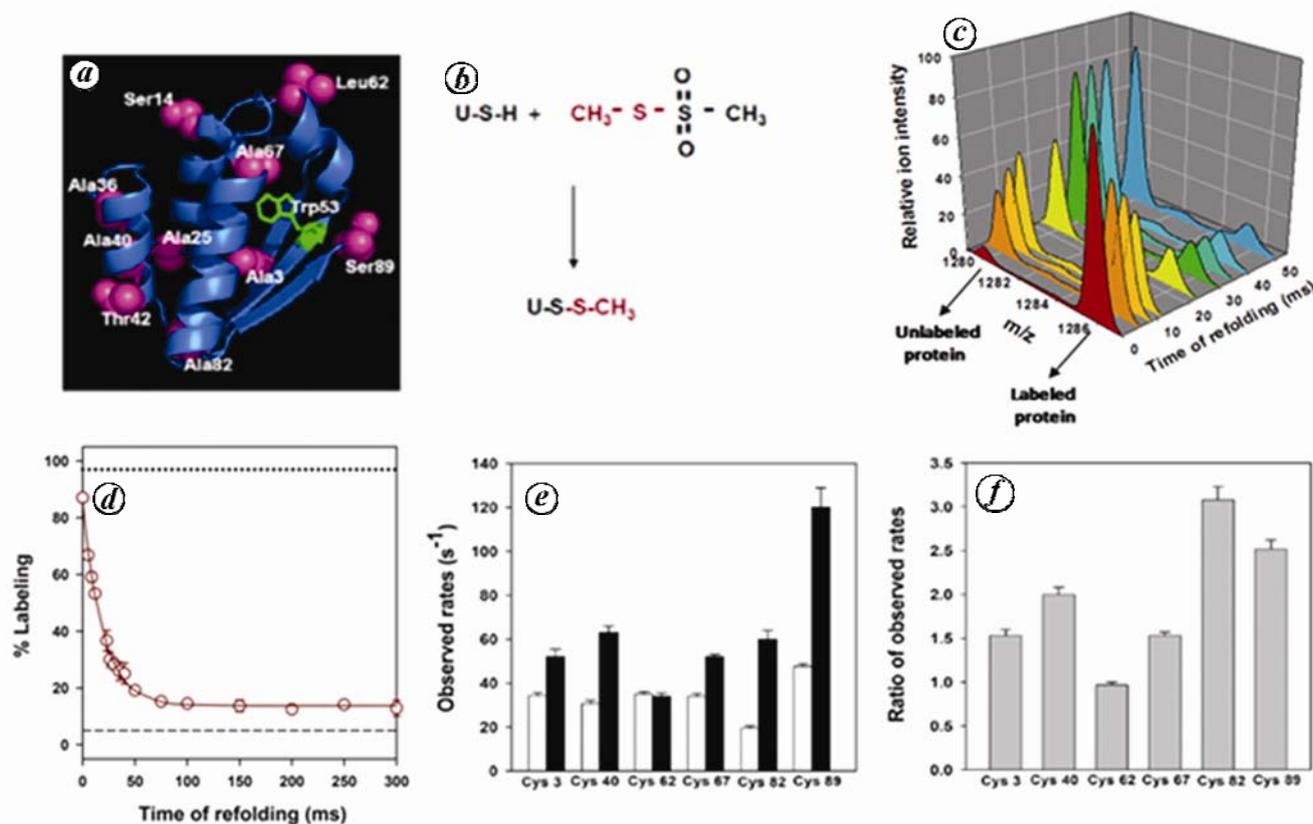
Figure 2. The major (ms) refolding reaction of barstar was studied using the pulse-thiol labelling methodology in conjunction with mass-spectrometry. *a*, Locations of side chains in the protein structure that were mutated to cysteine one at a time. *b*, *c*, The sole thiol group in each protein was 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[165].

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[162]. The main folding reaction also appeared to occur in a non-cooperative manner[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 apo-myoglobin[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[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[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[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[163,164]. It was shown that native barstar can sample the fully unfolded conformation even in the absence of denaturant[164], and that competing pathways are available to the protein for unfolding[163].

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[167–170]. 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[171]. An intermediate was shown to be populated on the folding pathway of engrailed homeodomain[172]. 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[173]. It is important to note that earlier equilibrium and stopped flow kinetic studies had indicated that the folding of acyl-CoA binding protein could be described by a 'two-state' mechanism[16]. 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[51], that the initial collapse reaction is a gradual structural transition[52,62,63], and that surface expansion occurs independently of core solvation during unfolding[174]. Interestingly, it was also shown that the otherwise spectroscopically silent *cis–trans* proline isomerization reaction can be directly monitored by FRET measurements during unfolding[174]. 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[22,175–177]. 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[177–179]. 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[179].

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[137,180]. Energy landscape theory for protein folding predicts that intermediates are ensembles of structurally distinct forms[55,141], 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 $\geq 5\,k_BT$), 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 ~ 1–2 k_BT) (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[28,65,181–184].

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
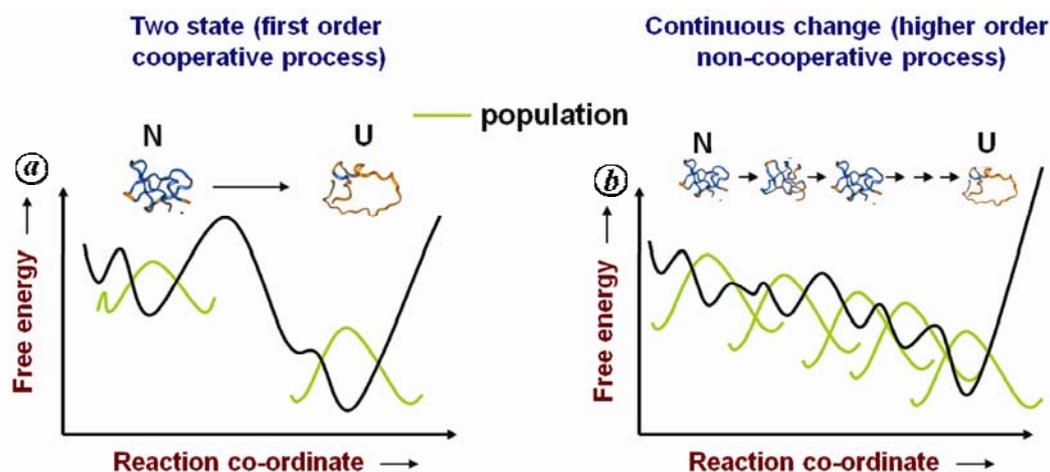
Figure 3. Energy diagrams for 'two-state' (*a*) and continuous (*b*) protein unfolding scenarios. In the 'two-state' unfolding scenario, one dominant free energy barrier between the N and U states ensures that only these two forms are populated either under different conditions or at different times of unfolding. In the alternate scenario, the unfolding reaction is mediated by a large number of small distributed barriers (~ 1–$2\ k_B T$). 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.

manner[27], and has demonstrated that parallel unfolding pathways operate during the unfolding of a variant of green fluorescent protein[185]. Recently, sm-FRET experiments have also been successful in revealing the conformational heterogeneity of the native structure, when it was shown that a mutant form of the protein Rop-homodimer can exist in two conformational sub-states under native-like conditions[186].

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 $\sim 200\ \mu s$ for the barrier crossing time, for a folding or unfolding reaction[187,188]. 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[189–191]. Single-molecule fluorescence studies of the slow unfolding reaction of green fluorescence protein have shown that each protein molecule jumps continuously between many conformational sub-states for many milliseconds, immediately before flipping to the U state[192,193]. 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)[194], 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[170]. 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[25,195–198], and in unfolded proteins[195,199,200]. Denaturant-dependent, non-random structure was shown to be present in the unfolded state of barstar[200]. Interestingly, in a separate study, it was shown that under conditions where ensemble-averaged probes suggested 'two-state' unfolding of barstar (Figure 4 *a*), TR-FRET measurements indicated that the structure
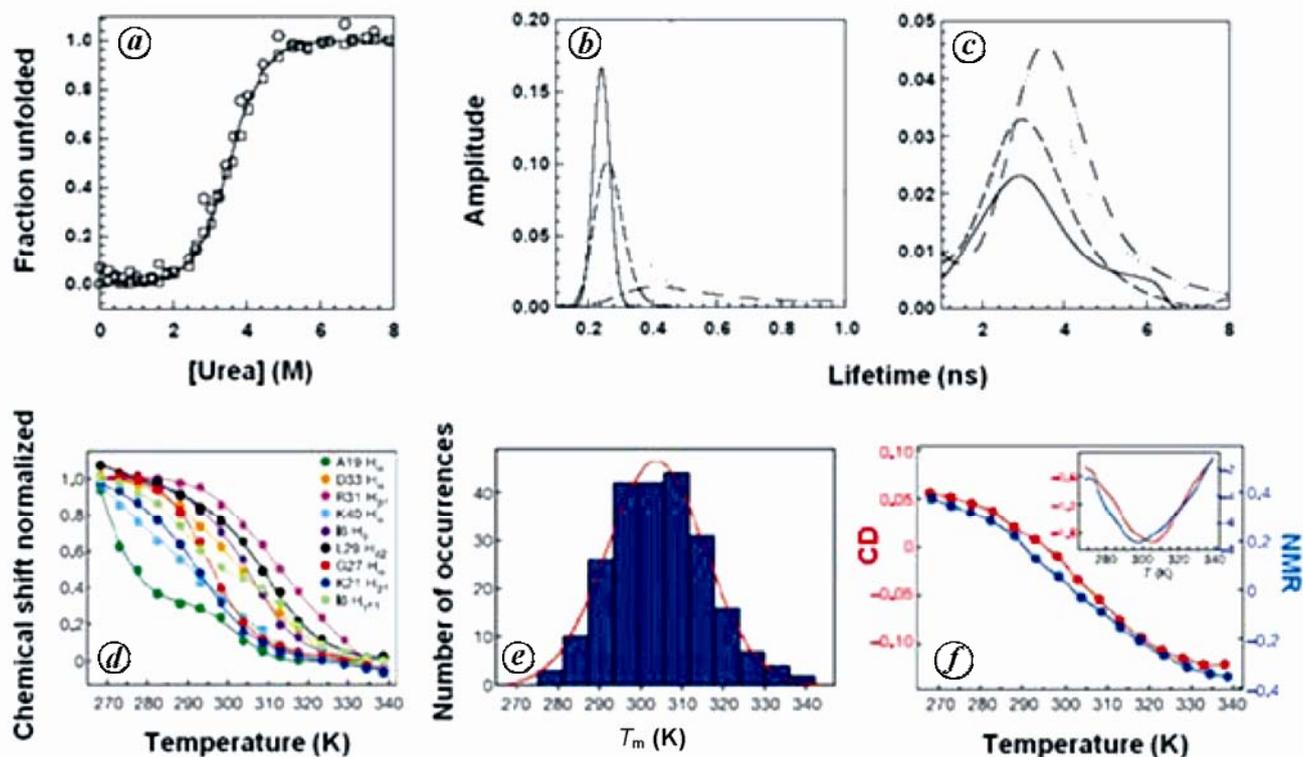
Figure 4. Structure is lost in a progressive manner during the unfolding of (*a–c*) barstar and (*d–f*) BBL. *a*, Change in the fraction of unfolded protein with different concentrations of urea as calculated from measurements of the fluorescence intensity at 360 nm (open circle) and the ellipticity 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[25] (Figure 4 *b* and *c*). TR-FRET measurements have also shown that conformational heterogeneity exists during the initial stages of the folding of cytochrome *c*[196], and TIM barrel protein[201].

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[198], as predicted by energy landscape theory[55]. 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[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[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 d). Also, there existed a wide distribution in the midpoints of the unfolding transitions monitored by different protons (Figure 4 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 f). Recently, the equilibrium unfolding of barstar was also studied using ^{19}F-NMR, and the data suggested that the protein unfolds in many steps[202] as had been indicated in a previous study[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[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[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[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[46]. Because of the non-specific nature of the hydrophobic interactions, alternative hydrophobic core-packing arrangements could exist in proteins[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[206], changing interaction partners[207] and

ligand binding[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[208].

The ability of proteins to fluctuate continuously between semi-stable states might be important for the acquisition of new traits during evolution[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[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[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[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[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[132,163,174,179,214–219] or the native state[157,160,164]. It has also been shown that even though a protein may unfold through multiple pathways[163,214], 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[220]. 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[220]. 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[221]. 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[222].

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[223]. 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[67,224–227]. The protein becomes unstable in denaturing conditions and overall unfolding occurs rapidly. The second hypothesis is the dry molten globule hypothesis[228–230], 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
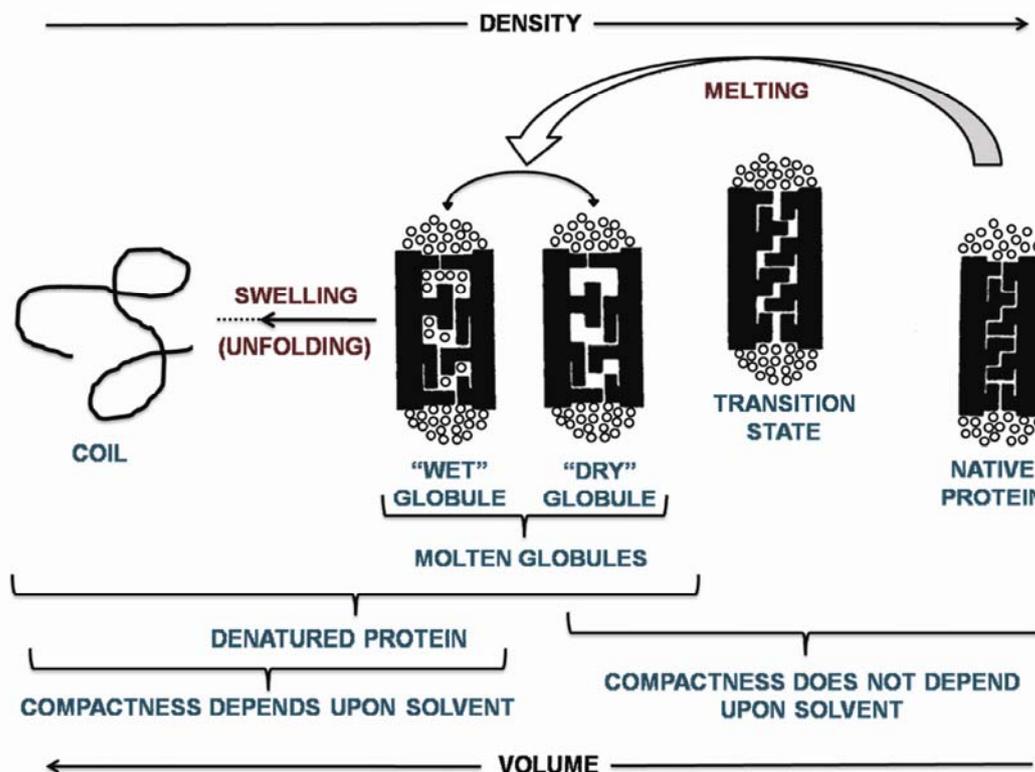

Figure 5. Scheme of protein denaturation based upon the dry molten globule hypothesis. Taken and modified with permission from Finkelstein and Shakhnovich[229] and Shakhnovich and Finkelstein[228].

measurements of the unfolding of ribonuclease A[223]. 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[227]. Similar results were obtained in studies of the unfolding of 6-^{19}F-tryptophan-labelled dihydrofolate reductase[231]. 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. ^{17}O relaxation dispersion NMR experiments have also helped in the identification of equilibrium dry molten globules for three proteins[232]. 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[178].

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[132,171,174]. 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, with substantial secondary structure content, to the random coil occurs in a gradual manner. Recently, the unfolding reaction of the small plant protein monellin was studied using a battery of biophysical techniques. These included changes in tryptophan and ANS (8-anilino-1-naphthalenesulfonic acid) fluorescence, near- and far-UV CD, as well as residue-specific probes such as multi-site steady-state FRET[233]. Far- and near-UV CD measurements of GdnHCl-induced unfolding indicated that a molten globule intermediate forms initially, before the major slow unfolding reaction commences. Steady-state FRET measurements showed that the C-terminal end of the single helix of monellin initially moves rapidly away from the single tryptophan residue that is close to the N-terminal end of the helix. The average end-to-end distance of the protein also expands during unfolding to the dry molten globule intermediate. This occurs without the entry of water molecules into the protein core, according to the evidence from intrinsic tryptophan fluorescence and ANS fluorescence-monitored kinetic unfolding measurements. Hence, these results provided direct evidence that the unfolding of monellin begins with an initial expansion of the protein into a dry

molten globule state, in which the sole helix has moved out of its place in the native structure[233].

In a separate but related study, the main unfolding reaction of monellin was probed further by measurement of the changes in the distributions of four different intramolecular distances, using a multi-site, TR-FRET methodology[234]. Interestingly, two out of the four distances measured were seen to expand in a gradual manner during unfolding, indicating that the protein molecules undergo slow and continuous, diffusive swelling and that specific structure is lost during the swelling process gradually, as predicted by the dry molten globule hypothesis (see above). The swelling process could be adequately modelled as the dynamics of a Rouse-like polymer chain, and the study brought out the polymer nature of protein folding and unfolding reactions. Surprisingly, the expansion of the other two distances appeared to occur in a seemingly 'two-state', all-or-none manner. These results highlight the importance of using multiple residue-specific probes to study protein folding or unfolding reactions, and indicate that the structural transition between the native and unfolded states can be a combination of first-order and higher-order transitions. This study further showed that the occurrence of single exponential kinetics does not necessarily indicate that the structural evolution is not gradual[235].

It is commonly believed that hydrophobic interactions are of paramount importance in determining the stability of a protein fold[67,224–226,236]. 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[46,237]. 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[238].

Concluding remarks

It is commonly believed that proteins fold and unfold in a cooperative, all-or-none manner over an energy landscape describable by a single dominant barrier on a single reaction coordinate. It is, however, becoming increasingly apparent, with the application of high-resolution probes, that the folding and unfolding transitions of proteins may be so highly non-cooperative as to occur in multiple steps or even gradually, so that many small free energy barriers distributed on a complex free energy landscape have to be crossed. The observation that some proteins can indeed fold or unfold through a continuum of intermediate forms on the timescale of several seconds is important. It indi-

cates 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.

1. Mirsky, A. E. and Pauling, L., On the structure of native, denatured, and coagulated proteins. *Proc. Natl. Acad. Sci. USA*, 1936, **22**, 439–447.

2. Anson, M. L., The denaturation of proteins by detergents and bile salts. *Science*, 1939, **90**, 256–257.

3. Levinthal, C., Are there pathways for protein folding? *J. Chim. Phys.*, 1968, **65**, 44–45.

4. Levinthal, C., How to fold graciously? In *Mossbauer Spectroscopy in Biological Systems: Proceedings of a meeting Held at Allerton House, Monticello, Illinois* (eds DeBrunner, J. T. P. and Munck, E.), University of Illinois Press, Urbana, 1969, pp. 22–24.

5. Anfinsen, C. B., Principles that govern the folding of protein chains. *Science*, 1973, **181**, 223–230.

6. Monod, J., *Chance and Necessity: An Essay on the Natural Philosophy of Modern Biology* (translated in English by Wainhouse, A.), Knopf, New York, 1971, pp. 95–96.

7. Taylor, W. R., May, A. C. W., Brown, N. P. and Aszódi, A., Protein structure: geometry, topology and classification. *Rep. Prog. Phys.*, 2001, **64**, 517–590.

8. Zwanzig, R., Szabo, A. and Bagchi, B., Levinthal's paradox. *Proc. Natl. Acad. Sci. USA*, 1992, **89**, 20–22.

9. Chiti, F. and Dobson, C. M., Amyloid formation by globular proteins under native conditions. *Nat. Chem. Biol.*, 2009, **5**, 15–22.

10. O'Brien, P. J. and Herschlag, D., Catalytic promiscuity and the evolution of new enzymatic activities. *Chem. Biol.*, 1999, **6**, R91–R105.

11. Wang, X., Minasov, G. and Shoichet, B. K., Evolution of an antibiotic resistance enzyme constrained by stability and activity trade-offs. *J. Mol. Biol.*, 2002, **320**, 85–89.

12. Murphy, P. M., Bolduc, J. M., Gallaher, J. L., Stoddard, B. L. and Baker, D., Alteration of enzyme specificity by computational loop remodeling and design. *Proc. Natl. Acad. Sci. USA*, 2009, **106**, 9215–9220.

13. Tracewell, C. A. and Arnold, F. H., Directed enzyme evolution: climbing fitness peaks one amino acid at a time. *Curr. Opin. Chem. Biol.*, 2009, **13**, 3–9.

14. Bloom, J. D. and Arnold, F. H., In the light of directed evolution: pathways of adaptive protein evolution. *Proc. Natl. Acad. Sci. USA*, 2009, **106**, 9995–10000.

15. Jackson, S. E. and Fersht, A. R., Folding of chymotrypsin inhibitor-2. 1. Evidence for a two-state transition. *Biochemistry*, 1991, **30**, 10428–10435.

16. Kragelund, B. B., Robinson, C. V., Knudsen, J., Dobson, C. M. and Poulsen, F. M., Folding of a four-helix bundle: studies of acyl-coenzyme A binding protein. *Biochemistry*, 1995, **34**, 7217–7224.

17. Jackson, S. E., How do small single-domain proteins fold? *Fold. Des.*, 1998, **3**, R81–R91.

18. Go, N., Theory of reversible denaturation of globular proteins. *Int. J. Pept. Protein Res.*, 1975, **7**, 313–323.

19. Go, N., Protein folding as a stochastic process. *J. Stat. Phys.*, 1983, **30**, 413–423.

20. Zocchi, G., Proteins unfold in steps. *Proc. Natl. Acad. Sci. USA*, 1997, **94**, 10647–10651.

21. Englander, S. W., Protein folding intermediates and pathways studied by hydrogen exchange. *Annu. Rev. Biophys. Biomol. Struct.*, 2000, **29**, 213–238.

22. Swaminathan, R., Nath, U., Udgaonkar, J. B., Periasamy, N. and Krishnamoorthy, G., Motional dynamics of a buried tryptophan reveals the presence of partially structured forms during denaturation of barstar. *Biochemistry*, 1996, **35**, 9150–9157.

23. Holtzer, M. E., Lovett, E. G., d'Avignon, D. A. and Holtzer, A., Thermal unfolding in a GCN4-like leucine zipper: ^{13}C$^\alpha$ NMR chemical shifts and local unfolding curves. *Biophys. J.*, 1997, **73**, 1031–1041.

24. Song, J., Jamin, N., Gilquin, B., Vita, C. and Menez, A., A gradual disruption of tight side-chain packing: 2D ^{1}H-NMR characterization of acid-induced unfolding of CHABII. *Nature Struct. Biol.*, 1999, **6**, 129–134.

25. Lakshmikanth, G. S., Sridevi, K., Krishnamoorthy, G. and Udgaonkar, J. B., Structure is lost incrementally during the unfolding of barstar. *Nature Struct. Biol.*, 2001, **8**, 799–804.

26. Ahmed, Z., Beta, I. A., Mikhonin, A. V. and Asher, S. A., UV-resonance Raman thermal unfolding study of Trp-cage shows that it is not a simple two-state miniprotein. *J. Am. Chem. Soc.*, 2005, **127**, 10943–10950.

27. Kuzmenkina, E. V., Heyes, C. D. and Nienhaus, G. U., Single-molecule FRET study of denaturant induced unfolding of RNase H. *J. Mol. Biol.*, 2006, **357**, 313–324.

28. Sadqi, M., Fushman, D. and Munoz, V., Atom-by-atom analysis of global downhill protein folding. *Nature*, 2006, **442**, 317–321.

29. Sinha, K. K. and Udgaonkar, J. B., Early events in protein folding. *Curr. Sci.*, 2009, **96**, 1053–1070.

30. Ptitsyn, O. B., Stages in the mechanism of self-organization of protein molecules. *Dokl. Akad. Nauk SSSR*, 1973, **210**, 1213–1215.

31. Kim, P. S. and Baldwin, R. L., Specific intermediates in the folding reactions of small proteins and the mechanism of protein folding. *Annu. Rev. Biochem.*, 1982, **51**, 459–489.

32. Udgaonkar, J. B. and Baldwin, R. L., NMR evidence for an early framework intermediate on the folding pathway of ribonuclease A. *Nature*, 1988, **335**, 694–699.

33. Karplus, M., and Weaver, D. L., Protein-folding dynamics. *Nature*, 1976, **260**, 404–406.

34. Bashford, D., Cohen, F. E., Karplus, M., Kuntz, I. D. and Weaver, D. L., Diffusion–collision model for the folding kinetics of myoglobin. *Proteins*, 1988, **4**, 211–227.

35. Karplus, M. and Weaver, D. L., Protein folding dynamics: the diffusion–collision model and experimental data. *Protein Sci.*, 1994, **3**, 650–668.

36. Baldwin, R. L. and Rose, G. D., Is protein folding hierarchic? I. Local structure and peptide folding. *Trends Biochem. Sci.*, 1999, **24**, 26–33.

37. Baldwin, R. L. and Rose, G. D., Is protein folding hierarchic? II. Folding intermediates and transition states. *Trends Biochem. Sci.*, 1999, **24**, 77–83.

38. Wetlaufer, D. B., Nucleation, rapid folding and globular intrachain regions in proteins. *Proc. Natl. Acad. Sci. USA*, 1973, **70**, 697–701.

39. Itzhaki, L. S., Otzen, D. E. and Fersht, A. R., The structure of the transition state for folding of chymotrypsin inhibitor 2 analyzed by protein engineering methods: evidence for a nucleation–condensation mechanism for protein folding. *J. Mol. Biol.*, 1995, **254**, 260–288.

40. Fersht, A. R., Optimization of rates of protein folding: the nucleation–condensation mechanism and its implications. *Proc. Natl. Acad. Sci. USA*, 1995, **92**, 10869–10873.

41. Fersht, A. R., Nucleation mechanisms in protein folding. *Curr. Opin. Struct. Biol.*, 1997, **7**, 3–9.

42. Daggett, V. and Fersht, A. R., Is there a unifying mechanism for protein folding? *Trends Biochem. Sci.*, 2003, **28**, 18–25.

43. Robson, B. and Pain, R. H., Analysis of the code relating sequence to conformation in proteins: possible implications for the mechanism of formation of helical regions. *J. Mol. Biol.*, 1971, **58**, 237–259.

44. Dill, K. A., Theory for the folding and stability of globular proteins. *Biochemistry*, 1985, **24**, 1501–1509.

45. Dill, K. A., Alonso, D. O. and Hutchinson, K., Thermal stabilities of globular protein. *Biochemistry*, 1989, **28**, 5439–5449.

46. Dill, K. A., Dominant forces in protein folding. *Biochemistry*, 1990, **29**, 7133–7155.

47. Alonso, D. O. and Dill, K. A., Solvent denaturation and stabilization of globular proteins. *Biochemistry*, 1991, **30**, 5974–5985.

48. Karplus, M. and Shakhnovich, E. I., Protein folding: theoretical studies of thermodynamics and dynamics. In *Protein Folding* (ed. Creighton, T.), W. H. Freeman and Co, New York, 1992.

49. Gutin, A. M., Abkevich, V. I. and Shakhnovich, E. I., Is burst hydrophobic collapse necessary for protein folding? *Biochemistry*, 1995, **34**, 3066–3076.

50. Dill, K. A., Bromberg, S., Yue, K., Fiebig, K. M., Yee, D. P., Thomas, P. D. and Chan, H. S., Principles of protein folding – a perspective from simple exact models. *Protein Sci.*, 1995, **4**, 561–602.

51. Agashe, V. R., Shastry, M. C. and Udgaonkar, J. B., Initial hydrophobic collapse in the folding of barstar. *Nature*, 1995, **377**, 754–757.

52. Sinha, K. K. and Udgaonkar, J. B., Dissecting the non-specific and specific components of the initial folding reaction of barstar by multi-site FRET measurements. *J. Mol. Biol.*, 2007, **370**, 385–405.

53. Bryngelson, J. D. and Wolynes, P. G., Spin glasses and the statistical mechanics of protein folding. *Proc. Natl. Acad. Sci. USA*, 1987, **84**, 7524–7528.

54. Bryngelson, J. D., Onuchic, J. N., Socci, N. D. and Wolynes, P. G., Funnels, pathways, and the energy landscape of protein folding: a synthesis. *Proteins*, 1995, **21**, 167–195.

55. Dill, K. A. and Chan, H. S., From Levinthal to pathways to funnels. *Nature Struct. Biol.*, 1997, **4**, 10–19.

56. Pande, V. S., Grosberg, A., Tanaka, T. and Rokhsar, D. S., Pathways for protein folding: is a new view needed? *Curr. Opin. Struct. Biol.*, 1998, **8**, 68–79.

57. Dill, K. A., The stabilities of globular proteins. In *Protein Engineering* (eds Oxender, D. L. and Fox, C. F.), A. R. Liss, New York, 1987, pp. 187–192.

58. Guo, Z. and Thirumalai, D., The nucleation–collapse mechanism in protein folding: evidence for the non-uniqueness of the folding nucleus. *Fold. Des.*, 1997, **2**, 377–391.

59. Miranker, A. D. and Dobson, C. M., Collapse and cooperativity in protein folding. *Curr. Opin. Struct. Biol.*, 1996, **6**, 31–42.

60. Eaton, W. A., Searching for 'downhill scenarios' in protein folding. *Proc. Natl. Acad. Sci. USA*, 1999, **96**, 5897–5899.

61. Gruebele, M., Downhill protein folding: evolution meets physics. *C. R. Biol.*, 2005, **328**, 701–712.

62. Sinha, K. K. and Udgaonkar, J. B., Dependence of the size of the initially collapsed form during the refolding of barstar on denaturant concentration: evidence for a continuous transition. *J. Mol. Biol.*, 2005, **353**, 704–718.

63. Sinha, K. K. and Udgaonkar, J. B., Barrierless evolution of structure during the sub-millisecond refolding reaction of a small protein. *Proc. Natl. Acad. Sci. USA*, 2008, **105**, 7998–8003.

64. Baldwin, R. L., The nature of protein folding pathways: the classical versus the new view. *J. Biomol. NMR*, 1995, **5**, 103–109.

65. Garcia-Mira, M. M., Sadqi, M., Fischer, N., Sanchez-Ruiz, J. M. and Muñoz, V., Experimental identification of downhill protein folding. *Science*, 2002, **298**, 2191–2195.

66. Ikai, A. and Tanford, C., Kinetics of unfolding and refolding of proteins. I. Mathematical analysis. *J. Mol. Biol.*, 1973, **73**, 145–163.

67. Kauzmann, W., Some factors in the interpretation of protein denaturation. *Adv. Protein Chem.*, 1959, **14**, 1–63.

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

69. Bolen, D. W. and Rose, G. D., Structure and energetics of the hydrogen-bonded backbone in protein folding. *Annu. Rev. Biochem.*, 2008, **77**, 339–362.

70. Matouschek, A., Kellis Jr, J. T., Serrano, L. and Fersht, A. R., Mapping the transition state and pathway of protein folding by protein engineering. *Nature*, 1989, **340**, 122–126.

71. Bycroft, M., Matouschek, A., Kellis Jr, J. T., Serrano, L. and Fersht, A. R., Detection and characterization of a folding intermediate in barnase by NMR. *Nature*, 1990, **346**, 488–490.

72. Karplus, M., Aspects of protein reaction dynamics: deviations from simple behavior. *J. Phys. Chem. B*, 2000, **104**, 11–27.

73. Chahine, J., Oliveira, R. J., Leite, V. B. and Wang, J., Configuration-dependent diffusion can shift the kinetic transition state and barrier height of protein folding. *Proc. Natl. Acad. Sci. USA*, 2007, **104**, 14646–14651.

74. Frauenfelder, H. and Wolynes, P. G., Rate theories and puzzles of heme protein kinetics. *Science*, 1985, **229**, 337–345.

75. Kramers, H. A., Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica (Utrecht)*, 1940, **7**, 284–304.

76. Hanggi, P., Talkner, P. and Borkovec, M., Reaction rate theory: fifty years after Kramers. *Rev. Mod. Phys.*, 1990, **62**, 251–341.

77. Shakhnovich, E. I., Theoretical studies of protein-folding thermo-dynamics and kinetics. *Curr. Opin. Struct. Biol.*, 1997, **7**, 29–40.

78. Socci, N. D., Onuchic, J. N. and Wolynes, P. G., Diffusive dynamics of the reaction coordinate for protein folding funnels. *J. Chem. Phys.*, 1996, **104**, 5860–5868.

79. Beece, D. *et al.*, Solvent viscosity and protein dynamics. *Biochemistry*, 1980, **19**, 5147–5157.

80. Brooks III, C. L., Karplus, M. and Pettitt, B. M., Proteins: a theoretical perspective of dynamics, structure and thermodynamics. *Adv. Chem. Phys.*, 1988, **71**, 1–249.

81. Ansari, A., Jones, C. M., Henry, E. R., Hofrichter, J. and Eaton, W. A., The role of solvent viscosity in the dynamics of protein conformational changes. *Science*, 1992, **256**, 1796–1798.

82. Bahar, I., Erman, B., Haliloglu, T. and Jernigan, R. L., Efficient characterization of collective motions and inter-residue correlations in proteins by low-resolution simulations. *Biochemistry*, 1997, **4**, 13512–13523.

83. Bahar, I. and Jernigan, R. L., Cooperative fluctuations and subunit communication in tryptophan synthase. *Biochemistry*, 1999, **38**, 3478–3490.

84. Amadei, A., de Groot, B. L., Ceruso, M.-A., Paci, M., Di Nola, A. and Berendsen, H. J. C., A kinetic model for the internal motions of proteins: diffusion between multiple harmonic wells. *Proteins Struct. Funct. Genet.*, 1999, **35**, 283–292.

85. Chrunyk, B. A. and Matthews, C. R., Role of diffusion in the folding of the α-subunit of tryptophan synthase from *Escherichia coli*. *Biochemistry*, 1990, **29**, 2149–2154.

86. Jacob, M., Schindler, T., Balbach, J. and Schmid, F. X., Diffusion control in an elementary protein folding reaction. *Proc. Natl. Acad. Sci. USA*, 1997, **94**, 5622–5627.

87. Creighton, T. E., Protein folding: does diffusion determine the folding rate? *Curr. Biol.*, 1997, **7**, R380–R383.

88. Plaxco, K. W. and Baker, D., Limited internal friction in the rate-limiting step of a two-state protein folding reaction. *Proc. Natl. Acad. Sci. USA*, 1998, **95**, 13591–13596.

89. Jacob, M. and Schmid, F. X., Protein folding as a diffusional process. *Biochemistry*, 1999, **38**, 13773–13779.

90. Jacob, M., Geeves, M., Holtermann, G. and Schmid, F. X., Diffusional barrier crossing in a two-state protein folding reaction. *Nature Struct. Biol.*, 1999, **6**, 923–926.

91. Pradeep, L. and Udgaonkar, J. B., Diffusional barrier in the unfolding of a small protein. *J. Mol. Biol.*, 2007, **366**, 1016–1028.

92. Ramos, C. H., Weisbuch, S. and Jamin, M., Diffusive motions control the folding and unfolding kinetics of the apomyoglobin pH 4 molten globule intermediate. *Biochemistry*, 2007, **46**, 4379–4389.

93. Cellmer, T., Henry, E. R., Hofrichter, J. and Eaton, W. A., Measuring internal friction of an ultrafast-folding protein. *Proc. Natl. Acad. Sci. USA*, 2008, **105**, 18320–18325.

94. Wensley, B. G. *et al.*, Experimental evidence for a frustrated energy landscape in a three-helix-bundle protein family. *Nature*, 2010, **463**, 685–688.

95. Leffler, J. E., Parameters for the description of transition states. *Science*, 1953, **117**, 340–341.

96. Hammond, G. S., A correlation of reaction rates. Hammond postulate. *J. Am. Chem. Soc.*, 1955, **77**, 334–338.

97. Jencks, W. P., A primer for the Bema Hapothle. An empirical approach to the characterization of changing transition-state structures. *Chem. Rev.*, 1985, **85**, 511–527.

98. Beasty, A. M., Hurle, M. R., Manz, J. T., Stackhouse, T., Onufer, J. J. and Matthews, C. R., Effects of the phenylalanine-22 → leucine, glutamic acid-49 → methionine, glycine-234 → aspartic acid, and glycine-234 → lysine mutations on the folding and stability of the α-subunit of tryptophan synthase from *Escherichia coli. Biochemistry*, 1986, **25**, 2965–2974.

99. Matthews, C. R., Effect of point mutations on the folding of globular proteins. *Methods Enzymol.*, 1987, **154**, 498–511.

100. Matthews, C. R. and Hurle, M. R., Mutant sequences as probes of protein folding mechanisms. *Bioessays*, 1987, **6**, 254–257.

101. Fersht, A. R. and Daggett, V., Protein folding and unfolding at atomic resolution. *Cell*, 2002, **108**, 573–582.

102. Fersht, A. R. and Sato, S., ϕ-Value analysis and the nature of protein-folding transition states. *Proc. Natl. Acad. Sci. USA*, 2004, **101**, 7976–7981.

103. Raleigh, D. P. and Plaxco, K. W., The protein folding transition state: what are ϕ-values really telling us? *Protein Pept. Lett.*, 2005, **12**, 117–122.

104. Milla, M. E., Brown, B. M., Waldburger, C. D. and Sauer, R. T., P22 Arc repressor: transition state properties inferred from mutational effects on the rates of protein unfolding and refolding. *Biochemistry*, 1995, **34**, 13914–13919.

105. Grantcharova, V. P., Riddle, D. S., Santiago, J. V. and Baker, D., Important role of hydrogen bonds in the structurally polarized transition state for folding of the src SH3 domain. *Nature Struct. Biol.*, 1998, **5**, 714–720.

106. Martinez, J. C., Pisabarro, M. T. and Serrano, L., Obligatory steps in protein folding and the conformational diversity of the transition state. *Nature Struct. Biol.*, 1998, **5**, 721–729.

107. Kippen, A. D., Sancho, J. and Fersht, A. R., Folding of barnase in parts. *Biochemistry*, 1994, **33**, 3778–3786.

108. Garcia-Mira, M. M., Boehringer, D. and Schmid, F. X., The folding transition state of the cold shock protein is strongly polarized. *J. Mol. Biol.*, 2004, **339**, 555–569.

109. Onuchic, J. N., Socci, N. D., Luthey-Schulten, Z. and Wolynes, P. G., Protein folding funnels: the nature of the transition state ensemble. *Fold. Des.*, 1996, **1**, 441–450.

110. Waldburger, C. D., Jonsson, T. and Sauer, R. T., Barriers to protein folding: formation of buried polar interactions is a slow step in acquisition of structure. *Proc. Natl. Acad. Sci. USA*, 1996, **93**, 2629–2634.

111. Burton, R. E., Myers, J. K. and Oas, T. G., Protein folding dynamics: quantitative comparison between theory and experiment. *Biochemistry*, 1998, **37**, 5337–5343.

112. Burns, L. L., Dalessio, P. M. and Ropson, I. J., Folding mechanism of three structurally similar β-sheet proteins. *Proteins*, 1998, **33**, 107–118.

113. Munson, M., Anderson, K. S. and Regan, L., Speeding up protein folding: mutations that increase the rate at which Rop folds and unfolds by over four orders of magnitude. *Fold. Des.*, 1997, **2**, 77–87.

114. Plaxco, K. W., Simons, K. T. and Baker, D., Contact order, transition state placement and the refolding rates of single domain proteins. *J. Mol. Biol.*, 1998, **277**, 985–994.

115. Zhang, T., Bertelsen, E., Benvegnu, D. and Alber, T., Circular permutation of T4 lysozyme. *Biochemistry*, 1993, **32**, 12311–12318.

116. Viguera, A. R., Blanco, F. J. and Serrano, L., The order of secondary structure elements does not determine the structure of a protein but does affect its folding kinetics. *J. Mol. Biol.*, 1995, **247**, 670–681.

117. Garrett, J. B., Mullins, L. S. and Raushel, F. M., Are turns required for the folding of ribonuclease T1? *Protein Sci.*, 1996, **5**, 204–211.

118. Otzen, D. E. and Fersht, A. R., Folding of circular and permuted chymotrypsin inhibitor 2: retention of the folding nucleus. *Biochemistry*, 1998, **37**, 8139–8146.

119. Lindberg, M. O., Tångrot, J., Otzen, D. E., Dolgikh, D. A., Finkelstein, A. V. and Oliveberg, M., Folding of circular permutants with decreased contact order: general trend balanced by protein stability. *J. Mol. Biol.*, 2001, **314**, 891–900.

120. Miller, E. J., Fischer, K. F. and Marqusee, S., Experimental evaluation of topological parameters determining protein-folding rates. *Proc. Natl. Acad. Sci. USA*, 2002, **99**, 10359–10363.

121. Silow, M. and Oliveberg, M., High-energy channeling in protein folding. *Biochemistry*, 1997, **36**, 7633–7637.

122. Olofsson, M., Hansson, S., Hedberg, L., Logan, D. T. and Oliveberg, M., Folding of S6 structures with divergent amino acid composition: pathway flexibility within partly overlapping foldons. *J. Mol. Biol.*, 2007, **365**, 237–248.

123. Lindberg, M. O. and Oliveberg, M., Malleability of protein folding pathways: a simple reason for complex behavior. *Curr. Opin. Struct. Biol.*, 2007, **17**, 21–29.

124. Cho, J.-H., Sato, S. and Raleigh, D. P., Thermodynamics and kinetics of non-native interactions in protein folding: a single point mutant significantly stabilizes the N-terminal domain of L9 by modulating non-native interactions in the denatured state. *J. Mol. Biol.*, 2004, **338**, 827–837.

125. Pradeep, L. and Udgaonkar, J. B., Effect of salt on the urea-unfolded form of barstar probed by *m*-value measurements. *Biochemistry*, 2004, **43**, 11393–11402.

126. Cho, J.-H. and Raleigh, D. P., Electrostatic interactions in the denatured state and in the transition state for protein folding: effects of denatured state interactions on the analysis of transition state structure. *J. Mol. Biol.*, 2006, **359**, 1437–1446.

127. Matthews, C. R. and Westmoreland, D. G., Nuclear magnetic resonance studies of residual structure in thermally unfolded ribonuclease A. *Biochemistry*, 1975, **14**, 4532–4538.

128. Neri, D., Billeter, M., Wider, G. and Wüthrich, K., NMR determination of residual structure in a urea-denatured protein, the 434-repressor. *Science*, 1992, **257**, 1559–1563.

129. Wong, K. B., Freund, S. M. and Fersht, A. R., Cold denaturation of barstar: ^{1}H, ^{15}N and ^{13}C NMR assignment and characterization of residual structure. *J. Mol. Biol.*, 1996, **259**, 805–818.

130. Nolting, B., Golbik, R., Soler-Gonzalez, A. S. and Fersht, A. R., Circular dichroism of denatured barstar suggests residual structure. *Biochemistry*, 1997, **36**, 9899–9905.

131. Shortle, D., The expanded denatured state: an ensemble of conformations trapped in a locally encoded topological space. *Adv. Protein Chem.*, 2002, **62**, 1–23.

132. Juneja, J. and Udgaonkar, J. B., NMR studies of protein folding. *Curr. Sci.*, 2003, **84**, 157–172.

133. Anil, B., Li, Y., Cho, J.-H. and Raleigh, D. P., The unfolded state of NTL9 is compact in the absence of denaturant. *Biochemistry*, 2006, **45**, 10110–10116.

134. Sanchez, I. E. and Kiefhaber, T., Origin of unusual ϕ-values in protein folding: evidence against specific nucleation sites. *J. Mol. Biol.*, 2003, **334**, 1077–1085.

135. Goldenberg, D. P., Finding the right fold. *Nature Struct. Biol.*, 1999, **6**, 987–990.

136. Ozkan, S. B., Bahar, I. and Dill, K. A., Transition states and the meaning of ϕ-values in protein folding kinetics. *Nature Struct. Biol.*, 2001, **8**, 765–769.

137. Poland, D. C. and Scheraga, H. A., Statistical mechanics of non-covalent bonds in polyamino acids. IX. The two-state theory of protein denaturation. *Biopolymers*, 1965, **3**, 401–419.

138. Plotkin, S. S. and Wolynes, P. G., Non-Markovian configurational diffusion and reaction coordinates for protein folding. *Phys. Rev. Lett.*, 1998, **80**, 5015–5018.

139. Gfeller, D., De Los Rios, P., Caflisch, A. and Rao, F., Complex network analysis of free energy landscapes. *Proc. Natl. Acad. Sci. USA*, 2007, **104**, 1817–1822.

140. Chan, H. S. and Dill, K. A., Transition states and folding dynamics of proteins and heteropolymers. *J. Chem. Phys.*, 1994, **100**, 9238–9257.

141. Wolynes, P. G., Onuchic, J. N. and Thirumalai, D., Navigating the folding routes. *Science*, 1995, **267**, 1619–1620.

142. Brooks III, C. L., Gruebele, M., Onuchic, J. N. and Wolynes, P. G., Chemical physics of protein folding. *Proc. Natl. Acad. Sci. USA*, 1998, **95**, 11037–11038.

143. Shimada, J. and Shakhnovich, E. I., The ensemble folding kinetics of protein G from an all-atom Monte Carlo simulation. *Proc. Natl. Acad. Sci. USA*, 2002, **99**, 11175–11180.

144. Bhuyan, A. K. and Udgaonkar, J. B., Observation of multistate kinetics during the slow folding and unfolding of barstar. *Biochemistry*, 1999, **38**, 9158–9168.

145. Pradeep, L. and Udgaonkar, J. B., Osmolytes induce structure in an early intermediate on the folding pathway of barstar. *J. Biol. Chem.*, 2004, **279**, 40303–40313.

146. Radford, S. E., Dobson, C. M. and Evans, P. A., The folding of hen lysozyme involves partially structured intermediates and multiple pathways. *Nature*, 1992, **358**, 302–307.

147. Elove, G. A., Chaffotte, A. F., Roder, H. and Goldberg, M. E., Early steps in cytochrome *c* folding probed by time-resolved circular dichroism and fluorescence spectroscopy. *Biochemistry*, 1992, **31**, 6876–6883.

148. Ma, H. and Gruebele, M., Kinetics are probe-dependent during downhill folding of an engineered λ_{6-85} protein. *Proc. Natl. Acad. Sci. USA*, 2005, **102**, 2283–2287.

149. Austin, R. H., Beeson, K. W., Eisenstein, L., Frauenfelder, H. and Gunsalus, I. C., Dynamics of ligand binding to myoglobin. *Biochemistry*, 1975, **14**, 5355–5373.

150. Sabelko, J., Ervin, J. and Gruebele, M., Observation of strange kinetics in protein folding. *Proc. Natl. Acad. Sci. USA*, 1999, **96**, 6031–6036.

151. Roy, M. and Jennings, P. A., Real-time NMR kinetic studies provide global and residue-specific information on the non-cooperative unfolding of the β-trefoil protein, interleukin-1β. *J. Mol. Biol.*, 2003, **328**, 693–703.

152. Schanda, P., Forge, V. and Brutscher, B., Protein folding and unfolding studied at atomic resolution by fast two-dimensional NMR spectroscopy. *Proc. Natl. Acad. Sci. USA*, 2007, **104**, 11257–11262.

153. Roder, H., Elove, G. A. and Englander, S. W., Structural characterization of folding intermediates in cytochrome *c* by H-exchange labeling and proton NMR. *Nature*, 1988, **335**, 700–704.

154. Udgaonkar, J. B. and Baldwin, R. L., Early folding intermediate of ribonuclease A. *Proc. Natl. Acad. Sci. USA*, 1990, **87**, 8197–8201.

155. Lu, J. and Dahlquist, F. W., Detection and characterization of an early folding intermediate of T4 lysozyme using pulsed hydrogen exchange and two dimensional NMR. *Biochemistry*, 1992, **31**, 4749–4756.

156. Jacobs, M. D. and Fox, R. O., Staphylococcal nuclease folding intermediate characterized by hydrogen exchange and NMR spectroscopy. *Proc. Natl. Acad. Sci. USA*, 1994, **91**, 449–453.

157. Bai, Y., Sosnick, T. R., Mayne, L. and Englander, S. W., Protein folding intermediates: native-state hydrogen exchange. *Science*, 1995, **269**, 192–197.

158. Llinás, M., Gillespie, B., Dahlquist, F. W. and Marqusee, S., The energetics of T4 lysozyme reveal a hierarchy of conformations. *Nature Struct. Biol.*, 1999, **6**, 1072–1078.

159. Juneja, J. and Udgaonkar, J. B., Characterization of the unfolding of ribonuclease A by a pulsed hydrogen exchange study: evidence for competing pathways for unfolding. *Biochemistry*, 2002, **41**, 2641–2654.

160. Wani, A. H. and Udgaonkar, J. B., Native state dynamics drive the unfolding of the SH3 domain of PI3 kinase at high denaturant concentration. *Proc. Natl. Acad. Sci. USA*, 2009, **106**, 20711–20716.

161. Ballery, N., Desmadril, M., Minard, P. and Yon, J. M., Characterization of an intermediate in the folding pathway of phosphoglycerate kinase: chemical reactivity of genetically introduced cysteinyl residues during the folding process. *Biochemistry*, 1993, **32**, 708–714.

162. Ha, J. H. and Loh, S. N., Changes in side chain packing during apomyoglobin folding characterized by pulsed thiol-disulfide exchange. *Nature Struct. Biol.*, 1998, **5**, 730–737.

163. Ramachandran, S., Rami, B. R. and Udgaonkar, J. B., Measurements of cysteine reactivity during protein unfolding suggest the presence of competing pathways. *J. Mol. Biol.*, 2000, **297**, 733–745.

164. Sridevi, K. and Udgaonkar, J. B., Unfolding rates of barstar determined in native and low denaturant conditions indicate the presence of intermediates. *Biochemistry*, 2002, **41**, 1568–1578.

165. Jha, S. K. and Udgaonkar, J. B., Exploring the cooperativity of the fast folding reaction of a small protein using pulsed thiol labeling and mass spectrometry. *J. Biol. Chem.*, 2007, **282**, 37479–37491.

166. Feng, Z., Ha, J. H. and Loh, S. N., Identifying the site of initial tertiary structure disruption during apomyoglobin unfolding. *Biochemistry*, 1999, **38**, 14433–14439.

167. Stryer, L., Fluorescence energy transfer as a spectroscopic ruler. *Annu. Rev. Biochem.*, 1978, **47**, 819–846.

168. Selvin, P. R., The renaissance of fluorescence resonance energy transfer. *Nature Struct. Biol.*, 2000, **7**, 730–734.

169. Wu, P. and Brand, L., Resonance energy transfer: methods and applications. *Anal. Biochem.*, 1994, **218**, 1–13.

170. Lakowicz, J. R., Energy transfer. In *Principles of Fluorescence Spectroscopy*, Springer, New York, 2006, 3rd edn, pp. 443–472.

171. Lillo, M. P., Szpikowska, B. K., Mas, M. T., Sutin, J. D. and Beechem, J. M., Real-time measurement of multiple intramolecular distances during protein folding reactions: a multisite stopped-flow fluorescence energy-transfer study of yeast phosphoglycerate kinase. *Biochemistry*, 1997, **36**, 11273–11281.

172. Huang, F., Settanni, G. and Fersht, A. R., Fluorescence resonance energy transfer analysis of the folding pathway of engrailed homeodomain. *Protein Eng. Des. Sel.*, 2008, **21**, 131–146.

173. Teilum, K., Maki, K., Kragelund, B. B., Poulsen, F. M. and Roder, H., Early kinetic intermediate in the folding of acyl-CoA binding protein detected by fluorescence labeling and ultrarapid mixing. *Proc. Natl. Acad. Sci. USA*, 2002, **99**, 9807–9812.

174. Sridevi, K. and Udgaonkar, J. B., Surface expansion is independent of and occurs faster than core solvation during the unfolding of barstar. *Biochemistry*, 2003, **42**, 1551–1563.

175. Broos, J., Visser, A. J. W. G., Engbersen, J. F. J., Verboom, W., vanHoek, A. and Reinhoudt, D. N., Flexibility of enzymes suspended in organic solvents probed by time-resolved fluorescence anisotropy. Evidence that enzyme activity and enantioselectivity

are directly related to enzyme flexibility. *J. Am. Chem. Soc.*, 1995, **117**, 12657–12663.

176. Fa, M., Karolin, J., Aleshkov, S., Strandberg, L., Johansson, L. B. and Ny, T., Time-resolved polarized fluorescence spectroscopy studies of plasminogen activator inhibitor type 1: conformational changes of the reactive center upon interactions with target proteases, vitronectin and heparin. *Biochemistry*, 1995, **34**, 13833–13840.

177. Sridevi, K., Juneja, J., Bhuyan, A. K., Krishnamoorthy, G. and Udgaonkar, J. B., The slow folding reaction of barstar: the core tryptophan region attains tight packing before substantial secondary and tertiary structure formation and final compaction of the polypeptide chain. *J. Mol. Biol.*, 2000, **302**, 479–495.

178. Rami, B. R. and Udgaonkar, J. B., Mechanism of formation of a productive molten globule form of barstar. *Biochemistry*, 2002, **41**, 1710–1716.

179. Rami, B. R., Krishnamoorthy, G. and Udgaonkar, J. B., Dynamics of the core tryptophan during the formation of a productive molten globule intermediate of barstar. *Biochemistry*, 2003, **42**, 7986–8000.

180. Tanford, C., Isothermal unfolding of globular proteins in aqueous urea solutions. *J. Am. Chem. Soc.*, 1964, **86**, 2050–2059.

181. Ferguson, N., Schartau, P. J., Sharpe, T. D., Sato, S. and Fersht, A. R., One-state downhill versus conventional protein folding. *J. Mol. Biol.*, 2004, **344**, 295–301.

182. Ferguson, N., Sharpe, T. D., Johnson, C. M., Schartau, P. J. and Fersht, A. R., Structural biology: analysis of 'downhill' protein folding. *Nature*, 2007, **445**, E14–E15.

183. Li, P., Oliva, F. Y., Naganathan, A. N. and Muñoz, V., Dynamics of one-state downhill protein folding. *Proc. Natl. Acad. Sci. USA*, 2009, **106**, 103–108.

184. Huang, F., Ying, L. and Fersht, A. R., Direct observation of barrier-limited folding of BBL by single-molecule fluorescence resonance energy transfer. *Proc. Natl. Acad. Sci. USA*, 2009, **106**, 16239–16244.

185. Orte, A., Craggs, T. D., White, S. S., Jackson, S. E. and Klenerman, D., Evidence of an intermediate and parallel pathways in protein unfolding from single-molecule fluorescence. *J. Am. Chem. Soc.*, 2008, **130**, 7898–7907.

186. Gambin, Y. *et al.*, Direct single-molecule observation of a protein living in two opposed native structures. *Proc. Natl. Acad. Sci. USA*, 2009, **106**, 10153–10158.

187. Rhoades, E., Cohen, M., Schuler, B. and Haran, G., Two-state folding observed in individual protein molecules. *J. Am. Chem. Soc.*, 2004, **126**, 14686–14687.

188. Chung, H. S., Louis, J. M. and Eaton, W. A., Experimental determination of upper bound for transition path times in protein folding from single-molecule photon-by-photon trajectories. *Proc. Natl. Acad. Sci. USA*, 2009, **106**, 11837–11844.

189. Rhoades, E., Gussakovsky, E. and Haran, G., Watching proteins fold one molecule at a time. *Proc. Natl. Acad. Sci. USA*, 2003, **100**, 3197–3202.

190. Xie, Z., Srividya, N., Sosnick, T. R., Pan, T. and Scherer, N. F., Single-molecule studies highlight conformational heterogeneity in the early folding steps of a large ribozyme. *Proc. Natl. Acad. Sci. USA*, 2004, **101**, 534–539.

191. Qu, X., Smith, G. J., Lee, K. T., Sosnick, T. R., Pan, T. and Scherer, N. F., Single-molecule nonequilibrium periodic Mg^{2+}-concentration jump experiments reveal details of the early folding pathways of a large RNA. *Proc. Natl. Acad. Sci. USA*, 2008, **105**, 6602–6607.

192. Baldini, G., Cannone, F. and Chirico, G., Pre-unfolding resonant oscillations of single green fluorescent protein molecules. *Science*, 2005, **309**, 1096–1100.

193. Baldini, G., Cannone, F., Chirico, G., Collini, M., Campanini, B., Bettati, S. and Mozzarelli, A., Evidence of discrete substates and unfolding pathways in green fluorescent protein. *Biophys. J.*, 2007, **92**, 1724–1731.

194. Huang, F., Lerner, E., Sato, S., Amir, D., Haas, E. and Fersht, A. R., Time-resolved fluorescence resonance energy transfer study shows a compact denatured state of the B domain of protein A. *Biochemistry*, 2009, **48**, 3468–3476.

195. Navon, A., Ittah, V., Landsman, P., Scheraga, H. A. and Haas, E., Distributions of intramolecular distances in the reduced and denatured states of bovine pancreatic ribonuclease A. Folding initiation structures in the C-terminal portions of the reduced protein. *Biochemistry*, 2001, **40**, 105–118.

196. Lyubovitsky, J. G., Gray, H. B. and Winkler, J. R., Mapping the cytochrome *c* folding landscape. *J. Am. Chem. Soc.*, 2002, **124**, 5481–5485.

197. Lee, J. C., Engman, K. C., Tezcan, F. A., Gray, H. B. and Winkler, J. R., Structural features of cytochrome *c'* folding intermediates revealed by fluorescence energy-transfer kinetics. *Proc. Natl. Acad. Sci. USA*, 2002, **99**, 14778–14782.

198. Sridevi, K., Lakshmikanth, G. S., Krishnamoorthy, G. and Udgaonkar, J. B., Increasing stability reduces conformational heterogeneity in a protein folding intermediate ensemble. *J. Mol. Biol.*, 2004, **337**, 699–711.

199. Haas, E., McWherter, C. A. and Scheraga, H. A., Conformational unfolding in the N-terminal region of ribonuclease A detected by nonradiative energy transfer: distribution of interresidue distances in the native, denatured, and reduced-denatured states. *Biopolymers*, 1988, **27**, 1–21.

200. Saxena, A. M., Udgaonkar, J. B. and Krishnamoorthy, G., Characterization of intra-molecular distances and site-specific dynamics in chemically unfolded barstar: evidence for denaturant-dependent non-random structure. *J. Mol. Biol.*, 2006, **359**, 174–189.

201. Wu, Y., Kondrashkina, E., Kayatekin, C., Matthews, C. R. and Bilsel, O., Microsecond acquisition of heterogeneous structure in the folding of a TIM barrel protein. *Proc. Natl. Acad. Sci. USA*, 2008, **105**, 13367–13372.

202. Li, H. and Frieden, C., Comparison of C40/82A and P27A C40/82A barstar mutants using ^{19}F-NMR. *Biochemistry*, 2007, **46**, 4337–4347.

203. Qiu, L., Pabit, S. A., Roitberg, A. E. and Hagen, S. J., Smaller and faster: the 20-residue Trp-cage protein folds in 4 μs. *J. Am. Chem. Soc.*, 2002, **124**, 12952–12953.

204. Fernandez, J. M. and Li, H., Force-clamp spectroscopy monitors the folding trajectory of a single protein. *Science*, 2004, **303**, 1674–1678.

205. Lim, W. A. and Sauer, R. T., Alternative packing arrangements in the hydrophobic core of λ repressor. *Nature*, 1989, **339**, 31–36.

206. Smock, R. G. and Gierasch, L. M., Sending signals dynamically. *Science*, 2009, **324**, 198–203.

207. Lee, G. M. and Craik, C. S., Trapping moving targets with small molecules. *Science*, 2009, **324**, 213–215.

208. Frauenfelder, H., Sligar, G. and Wolynes, P. G., The energy landscapes and motions of proteins. *Science*, 1991, **254**, 1598–1603.

209. Watters, A. L., Deka, P., Corrent, C., Callender, D., Varani, G., Sosnick, T. and Baker, D., The highly cooperative folding of small naturally occurring proteins is likely the result of natural selection. *Cell*, 2007, **128**, 613–624.

210. Sadqi, M., de Alba, E., Pérez-Jiménez, R., Sanchez-Ruiz, J. M. and Muñoz, V., A designed protein as experimental model of primordial folding. *Proc. Natl. Acad. Sci. USA*, 2009, **106**, 4127–4132.

211. Tokuriki, N. and Tawfik, D. S., Protein dynamism and evolvability. *Science*, 2009, **324**, 203–207.

212. Cho, S. S., Weinkam, P. and Wolynes, P. G., Origins of barriers and barrierless folding in BBL. *Proc. Natl. Acad. Sci. USA*, 2008, **105**, 118–123.

213. Calamai, M., Chiti, F. and Dobson, C. M., Amyloid fibril formation can proceed from different conformations of a partially unfolded protein. *Biophys. J.*, 2005, **89**, 4201–4210.

214. Zaidi, F. N., Nath, U. and Udgaonkar, J. B., Multiple intermediates and transition states during protein unfolding. *Nature Struct. Biol.*, 1997, **4**, 1016–1024.

215. Jamin, M., Yeh, S. R., Rousseau, D. L. and Baldwin, R. L., Sub-millisecond unfolding kinetics of apomyoglobin and its pH 4 intermediate. *J. Mol. Biol.*, 1999, **292**, 731–740.

216. 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.

217. Galani, D., Fersht, A. K. and Perrett, S., Folding of the yeast prion protein Ure2: kinetic evidence for folding and unfolding intermediates. *J. Mol. Biol.*, 2002, **315**, 213–227.

218. Wani, A. H. and Udgaonkar, J. B., HX–ESI–MS and optical studies of the unfolding of thioredoxin indicate stabilization of a partially unfolded, aggregation-competent intermediate at low pH. *Biochemistry*, 2006, **45**, 11226–11238.

219. Patra, A. K. and Udgaonkar, J. B., Characterization of the folding and unfolding reactions of single-chain monellin: evidence for multiple intermediates and competing pathways. *Biochemistry*, 2007, **46**, 11727–11743.

220. Wright, C. F., Lindorff-Larsen, K., Randles, L. G. and Clarke, J., Parallel protein-unfolding pathways revealed and mapped. *Nature Struct. Biol.*, 2003, **10**, 658–662.

221. Wani, A. H. and Udgaonkar, J. B., Revealing a concealed intermediate that forms after the rate-limiting step of refolding of the SH3 domain of PI3 kinase. *J. Mol. Biol.*, 2009, **387**, 348–362.

222. 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. Biol.*, 1998, **276**, 657–667.

223. Kiefhaber, T., Labhardt, A. M. and Baldwin, R. L., Direct NMR evidence for an intermediate preceding the rate-limiting step in the unfolding of ribonuclease A. *Nature*, 1995, **375**, 513–515.

224. Tanford, C., Protein denaturation. *Adv. Protein Chem.*, 1968, **23**, 121–242.

225. Tanford, C., Protein denaturation. C. Theoretical models for the mechanism of denaturation. *Adv. Protein Chem.*, 1970, **24**, 1–95.

226. Privalov, P. L., Stability of proteins: small globular proteins. *Adv. Protein Chem.*, 1979, **33**, 167–241.

227. 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.

228. Shakhnovich, E. I. and Finkelstein, A. V., Theory of cooperative transitions in protein molecules. I. Why denaturation of globular protein is a first-order phase transition. *Biopolymers*, 1989, **28**, 1667–1680.

229. Finkelstein, A. V. and Shakhnovich, E. I., Theory of cooperative transitions in protein molecules. II. Phase diagram for a protein molecule in solution. *Biopolymers*, 1989, **28**, 1681–1694.

230. Hua, L., Zhou, R., Thirumalai, D. and Berne, B. J., Urea denaturation by stronger dispersion interactions with proteins than water implies a 2-stage unfolding. *Proc. Natl. Acad. Sci. USA*, 2008, **105**, 16928–16933.

231. Hoeltzli, S. D. and Frieden, C., Stopped-flow NMR spectroscopy: real-time unfolding studies of 6-^{19}F-tryptophan-labeled *Escherichia coli* dihydrofolate reductase. *Proc. Natl. Acad. Sci. USA*, 1995, **92**, 9318–9322.

232. Denisov, V. P., Jonsson, B. H. and Halle, B., Hydration of denatured and molten globule proteins. *Nature Struct. Biol.*, 1999, **6**, 253–260.

233. Jha, S. K. and Udgaonkar, J. B., Direct evidence for a dry molten globule intermediate during the unfolding of a small protein. *Proc. Natl. Acad. Sci. USA*, 2009, **106**, 12289–12294.

234. Jha, S. K., Dhar, D., Krishnamoorthy, G. and Udgaonkar, J. B., Continuous dissolution of structure during the unfolding of a small protein. *Proc. Natl. Acad. Sci. USA*, 2009, **106**, 11113–11118.

235. Hagen, S. J., Exponential decay kinetics in 'downhill' protein folding. *Proteins*, 2003, **50**, 1–4.

236. Privalov, P. L. and Gill, S. J., Stability of protein structure and hydrophobic interactions. *Adv. Protein Chem.*, 1988, **39**, 191–234.

237. Privalov, P. L. and Makhatadze, G. I., Contribution of hydration to protein folding thermodynamics. II. The entropy and Gibbs energy of hydration. *J. Mol. Biol.*, 1993, **232**, 660–679.

238. Segawa, S. and Sugihara, M., Characterization of the transition state of lysozyme unfolding. I. Effect of protein–solvent interactions on the transition state. *Biopolymers*, 1984, **23**, 2473–2488.