Eur Biophys J (2008) 37:721–728
DOI 10.1007/s00249-007-0256-x

REVIEW

Mechanisms of protein folding

Ylva Ivarsson · Carlo Travaglini-Allocatelli · Maurizio Brunori · Stefano Gianni

Received: 19 September 2007 / Accepted: 17 December 2007 / Published online: 9 January 2008

Abstract Understanding the mechanism by which a polypeptide chain folds into its native structure is a central problem of modern biophysics. The collaborative efforts of experimental and theoretical studies recently raised the tantalizing possibility to define a unifying mechanism for protein folding. In this review we summarize some of these intriguing advances and analyze them together with a discussion on the new findings concerning the so-called downhill folding.

Keywords Protein folding · Transition state · Nucleation–condensation · Diffusion–collision · Folding kinetics · Downhill folding

Abbreviations

TS	Transition state
DC	Diffusion–collision
NC	Nucleation–condensation
En-HD	Engrailed homeodomain
CI2	Chymotrypsin inhibitor 2
VHP	Chicken villin headpiece

Introduction

Discovering the details of the protein folding reaction is one of the central problems of structural biology and protein biophysics. Christian B. Anfinsen and Cyrus Levinthal may be considered as the historical forefathers of in vitro protein folding studies. In fact, while Anfinsen first observed that all the information needed for a protein to fold may be found in its amino acidic sequence (Anfinsen et al. 1961), Levinthal pointed out that folding cannot take place via a random sampling of all possible conformations (Levinthal 1968). Thus, the traditional view of the protein folding reaction was based on the existence of a preferred route driving the denatured chain to its native conformation via a sequence of consecutive intermediates, whose structural features are encoded by the primary structure. Novel theoretical and experimental approaches have recently contributed to challenge this mechanism and introduced more complex and detailed models, which capture the finest details of the folding reaction (Daggett and Fersht 2003).

Despite the inherent complexity of the protein folding reactions, proteins seem to fold via a limited number of accumulated intermediates with several proteins displaying a simple two-state behavior. Thus, a key role in protein folding studies has been played by the description of folding transition states (Fersht 2000). The transition state (TS) theory was introduced by Eyring in 1935 to model the observed kinetics of simple chemical inter-conversions (Eyring 1935). Following this view, observed kinetics of chemical reactions are dependent upon the formation of an activated complex, which arises from the productive collision of reactants. The TS of any chemical reaction may be defined as that particular configuration, along the reaction coordinate, that corresponds to the highest free energy. In theory when and if a reactant explores the transition state configuration, it will have an identical probability to revert to the reactant or progress to the product of a given chemical reaction. Fersht and coworkers extended the TS theory to protein folding studies and, as described below, contributed to the characterization of classical folding mechanisms. In

Y. Ivarsson · C. Travaglini-Allocatelli · M. Brunori · S. Gianni (✉)
Istituto di Biologia e Patologia Molecolari del CNR,
Dipartimento di Scienze Biochimiche "A. Rossi Fanelli",
Università di Roma "La Sapienza",
Piazzale A. Moro 5, 00185 Rome, Italy
e-mail: stefano.gianni@uniroma1.it

analogy with the TS of chemical reactions, the transition state ensemble of folding is the ensemble of protein conformations with an equal probability to fold or unfold. By employing a procedure developed by Fersht and co-workers, the ϕ value analysis (Fersht et al. 1992), interactions formed in the TS of folding have been mapped out for several small proteins; this approach has provided information on both native and non-native contacts formed in the TS. By systematically perturbing native interactions via mutagenesis while measuring the folding kinetics of the mutant protein, it is possible to identify interaction patterns present in the folding transition state (Matouschek et al. 1990). The role of a given residue is measured by its Φ value upon mutagenesis, which normalizes the change in stability of the activation barrier to that of the native state. The Φ value is a structural index for a given residue in the transition states or folding intermediates; Φ value of 1 suggests that the probed position is fully structured in the transition state whereas Φ value of 0 indicates that it is fully unstructured. While the interpretation of experimentally determined Φ values is widely debated (Fersht and Sato 2004; Weikl and Dill 2007), the Φ value analysis is to date the most widely employed technique to characterize protein-folding pathways.

In this review we will address some of the recent work focused on the describing the general mechanisms of the protein folding reaction. First, we will present the classical one-trajectory mechanisms used to describe the folding process, and highlight cases where apparently contradicting models of protein folding may manifest themselves in the folding of one-and-the-same protein. Then we focus on the statistical landscape view of protein folding and its implications with parallel folding channels and the possibilities of tuning the folding process into different trajectories. Finally, we summarize the recent findings and controversies on so-called downhill folding.

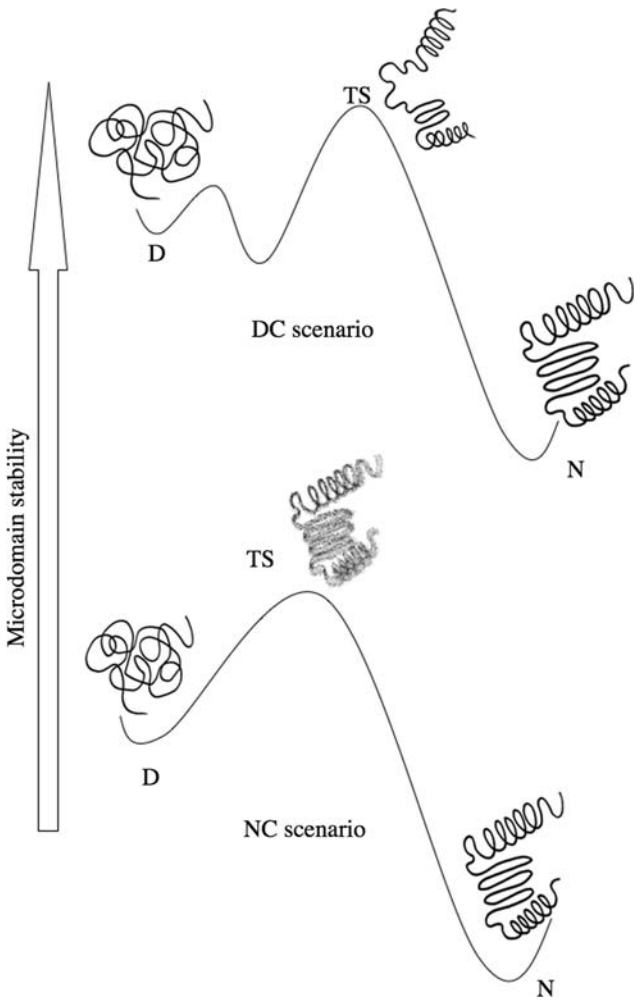

Fig. 1 Schematical representation of the classical mechanisms for protein folding described in the text. The diffusion–collision (*DC*) model (*top panel*) is based on the existence of micro-domains, involving segments of secondary structure, whose productive collision represents the rate-limiting step of folding. On the other hand, according to the nucleation–condensation (*NC*) model (*bottom panel*), the main folding TS is a distorted version of the native state and folding occurs via simultaneous formation of secondary and tertiary structure, around a weakly-structured folding nucleus

Classical mechanisms for the protein folding reaction

Two distinct mechanisms have classically been used to describe the folding of small single-domain proteins (Fig. 1). Some proteins, such as apomyoglobin (Bashford et al. 1988; Nishimura et al. 2000), barnase (Wong et al. 2000), the engrailed homeodomain (En-HD) (Mayor et al. 2003a) and the lambda repressor fragment (Myers and Oas 1999) have been described to fold in a stepwise manner, with rapid formation of distinct nuclei, comprising nascent secondary structure elements, followed by their collision and consolidation (diffusion–collision model, DC) (Karplus and Weaver 1994). Some other proteins, with chymotrypsin inhibitor 2 (CI2) as a first example (Itzhaki et al. 1995), seem to follow the nucleation–condensation model (NC) (Abkevich et al. 1994), with concurrent formation of secondary and tertiary structure. As discussed below, the magnitude and distribution of Φ values for a given protein is a critical test to distinguish between DC and NC mechanisms. In this section, we will summarize some of the characteristic features that, over the past decade, have been assigned to the DC and NC models.

In 1976 Karplus and Weaver discussed the dynamics of protein folding and introduced a plausible mechanism, the DC model (Karplus and Weaver 1976, 1994). This formalism assumes that a protein molecule is divided into parts such that the information contained in the amino acid sequence of each part is sufficient to fold autonomously. In particular, the DC model is based on the existence of

fluctuating quasi-particles, the "microdomains", which may be portions of incipient secondary structure. The microdomains may move diffusively and microdomain–microdomain collisions take place. Productive collisions lead to coalescence into intermediates, which may involve microdomains that are not necessarily contiguous along the protein sequence. Folding proceeds as a series of coalescence events that might either follow a unique order (sequential folding pathways) or explore different routes (parallel folding pathways), whose dominant components might be affected by solvent conditions or primary structure compositions for different homologous proteins. Thus, the DC model reduces the dynamics of the folding process to the diffusion events of discrete pre-formed microdomains of nascent secondary structure elements.

Two main corollaries are implied by the DC scenario (Karplus and Weaver 1994): (1) folding proceeds via a series of intermediates; and (2) formation of secondary structure elements precedes formation of tertiary structure. A paradigm example of a protein following the DC model is represented by the small helical protein En-HD. The folding pathway of En-HD involves the formation of a compact folding intermediate, whose structural properties have been addressed by protein engineering (Gianni et al. 2003b), small angle X-ray scattering (Mayor et al. 2003a, b) and NMR (Religa et al. 2005, 2007). Furthermore, in agreement with the DC model, the main folding transition state of En-HD displays heterogeneous structure localization, with regions having Φ values close to 1 and others displaying Φ values close to 0 distributed in contiguous blocks, indicative of preformed secondary structure elements (or independent microdomains) (Gianni et al. 2003a). In fact, recent investigation of the isolated helices 2 and 3 of the En-HD, forming a helix-turn-helix motif in the full domain, have demonstrated the helix-turn-helix motif to fold as an independent domain, and thus acting as an on-pathway intermediate (Religa et al. 2007).

The folding mechanism of CI2 has historically played a key role in our understanding of protein folding pathways. Indeed, the investigation of the folding pathway of CI2 by extensive Φ value analysis provided the opportunity to challenge the DC model and suggested the definition of the NC model. In 1991, CI2 was found as the first example of a two-state folding protein, where only the fully denatured and native states may be populated (Jackson and Fersht 1991), which clearly contradicted the DC model. Several other proteins were later found to follow the two-state mechanism (Jackson 1998). Furthermore, the structure of the folding transition state of CI2 was mapped by Φ value analysis on more than 100 mutants under various experimental conditions (Itzhaki et al. 1995); nearly all positions probed in the sequence of CI2 displayed fractional Φ values, with an average of ca. 0.3, indicating fractional bond formation of many of the probed interactions. Only a small set of positions, including a α-helix and a β-strand, displayed higher clustered Φ values, representing the "folding nucleus". From the overall fractional Φ values it appears as if the TS in the folding of CI2 is a distorted version of the native state and the protein seem to fold as a single cooperative unit, thus challenging the concept of pre-existing microdomains postulated by the DC model.

Consequently, the NC model postulates the existence of a folding nucleus, about which the transition state is constructed (Abkevich et al. 1994). Importantly, formation of the nucleus is not the solely rate-limiting event, and a significant portion of the structure must collapse (or condense) to an approximately correct conformation for the discontinuous network of interactions in the nucleus to come together. Formation of the nucleus (nucleation) is coupled with a more extended formation of structure (condensation) (Fersht 1995). Thus, the protein appears to collapse around an extended nucleus as this is being consolidated, and the TS appears as a distorted version of the native state with secondary and tertiary structure forming in parallel. By applying the NC model, the nucleus may be identified by the few residues displaying higher Φ values; the native-like structure in the TS should gradually decrease with a smooth gradient of decreasing Φ values from the nucleus, as exemplified by CI2 folding (Itzhaki et al. 1995). Interestingly, despite the weakness of the folding nucleus implied by the NC model, ingenious experiments based on topological mutations of CI2, including circular permutation (Otzen and Fersht 1997) and splitting of the protein (Neira et al. 1996), suggest that the position of the nucleus is quite conserved in proteins following the pure NC model.

Unifying features in protein folding mechanism

As exemplified above, kinetic experiments carried out in solution over the past two decades provided the basis for a major distinction between proteins folding with two- or three-state mechanisms, following the NC and DC models, respectively. Hence, in this distinction, the presence or absence of intermediates in the kinetic folding mechanism played a crucial role, and the methods of identifying their mechanistic roles were the topic of recent reviews (Brockwell and Radford 2007; Gianni et al. 2007c). Small proteins often fold fast and without detectable folding intermediates, which was taken as evidence against the importance of partially folded states in the folding of these proteins (Krantz et al. 2002). However, several studies indicate folding intermediates as consensus species even in the smallest and fastest folding proteins studied (Mayor et al. 2003b; Jemth et al. 2004).

Is there an inherent difference between two-state and multistate folders? To address this question, Sánchez and

Kiefhaber (2003) analyzed experimental data of 23 proteins, supposedly two-state folders, displaying deviations from linearity in their chevron plots which had previously been explained in terms of a broad TS barrier. According their analysis, the data were systematically fitted with great accuracy to a sequential model involving unstable, high energy intermediates, suggesting that the major difference between apparent two-state and multi-state folding proteins is the relative stability of the partially folded intermediates.

Studies of folding mechanism of homologous protein suggested that the structural mechanisms of folding may be conserved, even if the kinetic folding mechanism is different. The bacterial immunity proteins Im9 and Im7 apparently fold by two-state or three-state mechanisms, respectively (Ferguson et al. 1999). However, the late transition state ensembles of the two proteins have similar properties (Friel et al. 2003). In fact, under acidic conditions it was possible to stabilize a high-energy folding intermediate of Im9 so that it became kinetically detectable (Gorski et al. 2001). Furthermore, the folding intermediate of Im7 involves formation of non-native contacts (Capaldi et al. 2002). By rational engineering of Im9 it was possible to stabilize non-native interactions in the folding intermediate and thereby speed up the folding process (Friel et al. 2004). Similar switch in the folding mechanism has been observed in the case of RNase H (Spudich et al. 2004).

In the case of c-type cytochromes, analyses of the folding mechanism of cytochromes from prokaryotes to eukaryotes indicated that the folding pathway of these proteins involves the formation of one and the same partially structured intermediate (Travaglini-Allocatelli et al. 2004). In agreement with the proposed consensus mechanism, it was shown that the high-energy folding intermediate of *Pseudomonas aeruginosa* cyt *c*551 (Gianni et al. 2001) can be stabilized by rational mutagenesis so that it becomes kinetically detectable (Borgia et al. 2006).

As shown by these studies, there is no clear distinction between two-state folding through a NC model and three-state folders displaying DC behavior**.** Accordingly, recent work on the homeodomain superfamily (Gianni et al. 2003a) and on PDZ domains (Gianni et al. 2007c) suggested that NC and DC models may represent extreme manifestations of an underlying common mechanism and that proteins may appear to fold by either NC or DC depending on the inherent stability of their secondary structure elements. The homeodomain superfamily is composed of small three-helical bundle proteins. The kinetics and mechanisms of folding vary across the family, such that the human telomeric repeat binding factor 1, human repressor–activator protein 1 and the third repeat of mouse cellular myeoblastis DNA binding domain display two-state kinetics while the folding of En-HD displays three-state kinetics, as discussed previously (Gianni et al. 2003a). By different experimental and computational approaches, such as Φ value analysis and molecular dynamic simulations (White et al. 2005) it was shown that the mechanism of folding in the homedomain family changes from NC behavior to DC behavior, with helical elements fully preformed, concurrently with an increase in the inherent helical propensity. From these observations it was proposed a unifying mechanism that, depending on the inherent stability of the secondary structure elements, may manifest itself as either NC or DC behavior (Gianni et al. 2003a).

In agreement with the suggested unifying mechanism, the folding process of a PDZ domain was shown to comprise elements of both the DC and NC mechanisms (Gianni et al. 2007b). Folding of PDZ2 involves the crossing of two consecutive TS barriers, as shown by non-linear folding kinetics (Gianni et al. 2005) and ligand-induced folding experiments (Ivarsson et al. 2007). The interactions formed in the two distinct TS were mapped by ϕ value analysis. Surprisingly, while the first TS clearly displayed characteristics of folding according the NC mechanism (i.e. many mutants displaying $\phi = 0$ and only some mutants, corresponding to the folding nucleus, having fractional ϕ values), the second TS displayed characteristics of the DC mechanism (several ϕ close to 1, and the rest displaying fractional ϕ) (Gianni et al. 2007b). The folding of PDZ2 was suggested as a paradigmatic example in which the two extremes of the unifying mechanism are manifested in one single protein. According to this unifying folding mechanism, the folding of small globular proteins is suggested to involve three major events: (1) formation of a weak nucleus that determines the native-like topology of the structure, (2) a global collapse of the entire polypeptide chain, and finally, (3) consolidation of the remaining partially structured regions to achieve the native state conformations.

Landscape theory and multiple folding pathways

The folding of a protein involves formation and breakage of a myriad of contacts. Hence, the classical one-trajectory view of protein folding is an over-simplification of the mechanism of the folding process at a molecular level. A different view, the statistical energy landscape perspective (Onuchic et al. 1996, 1997), describes folding as taking place on a rugged free-energy landscape in which free-energy barriers separate ensembles of states displaying different levels of structural heterogeneity, such as the native and the denatured states. The high speed of protein folding, as compared to most barrier-limited chemical reactions, is due to the near-cancellation of the contributions of enthalpy and entropy to the free energy landscape during the folding process. The plasticity of folding pathways implicit in the landscape theory assumes that proteins can be rerouted through the energy landscape by mutational (Wright et al. 2003), topological (Lindberg and

Oliveberg 2007) or solvent perturbations (Otzen and Fersht 1997; Gianni et al. 2007c). The given trajectory taken by a protein molecule is determined by thermodynamic probabilities. In accordance with the landscape view of protein folding, the shift between NC and DC behavior highlights the tunability of the folding landscape.

In agreement with the statistical landscape theory, several proteins have been suggested to fold via parallel routes, such as lysozyme (Kiefhaber 1995) and c-type cytochromes (Gianni et al. 2003b). The first example of parallel folding pathways studied by Φ value analysis is the 27th immunoglobulin domain of titin (Wright et al. 2003). The compact TS of the predominant pathway was destabilized at increasing denaturant concentration and the flux of protein molecules were directed towards an alternative folding channel with less structured TS. Other experimental (Gianni et al. 2007a) and computational studies (Wallin et al. 2007) have reported proteins to switch their preferred folding pathways depending on the environmental conditions. The different TS observed in homologous proteins (McCallister et al. 2000) and in engineered circular permutants (Lindberg et al. 2002; Hubner et al. 2006) also support the notion of parallel folding pathways.

Molecular dynamic simulations have proposed the existence of parallel folding pathways for a variety of small protein domains (Lazaridis and Karplus 1997; Borreguero et al. 2004; Caflisch 2004; Rao et al. 2005; Juraszek and Bolhuis 2006; Lam et al. 2007) and interplay between experiments and simulations have surely contributed to the knowledge of the folding landscapes. However, as exemplified below, while the synergy of simulations and experiments contributed to broaden our knowledge on the finest details of the folding process (Daggett and Fersht 2003), care must be taken in over-interpreting results from molecular dynamic simulations without experimental validation. The symmetrical B domain of protein A (BdpA), a three-helix bundle protein of 60 residues, has been subjected to a number of molecular dynamic simulations (Wolynes 2004 and references therein). In a recent study (Itoh and Sasai 2006) this highly symmetrical protein was suggested to have access to multiple folding pathways. In the search of experimental evidence of parallel folding trajectories, Sato and Fersht (2007) performed an extensive ϕ value analysis of BdpA under various experimental conditions. To perturb the relative flow through different folding channels the folding experiments were performed at various temperatures and ϕ values were determined as a function of temperature in the range of 25–60°C. Despite significant change of folding conditions the ϕ values remain largely unchanged indicating the TS structure of BdpA remained more or less the same. Contrasting with the proposed parallel pathway scheme, the results rather suggest the folding of BpdA to take place through a single dominant pathway.

Downhill folding: when the barrier is just absent

Downhill folding, that is folding in the absence of barriers between the unfolded and native states, has lately received lot of attention. It is predicted by landscape theory that the energy barriers between the folded and the unfolded states are minimized when there is an extreme energetic bias towards the native state, such as at low temperature or in the absence of denaturant. Under such conditions, two-state folders or sequential intermediate models are predicted to be turned into downhill folding which proceeds through an array of temporary conformations with broad distributions of folding times (Bryngelson et al. 1995). The downhill folding is expected to take place very rapidly, approaching the speed limit of folding, which is about $N/100$ μs (where N is the number of residues) depending on the protein sequence and structure; for a full list of proteins folding in an ultra-fast time scale see Kubelka et al. (2004). Much of our current knowledge of the speed-limit for formation of simple elements of secondary structure has been achieved by systematically inspecting the folding of simple elements involving α-helix (Williams et al. 1996; Huang et al. 2001; Lapidus et al. 2002), β-hairpins (Munoz et al. 1997) and loops (Bieri et al. 1999; Buscaglia et al. 2005). For a three-helix bundles protein downhill folding would occur on timescales as short as about 0.6 μs.

Besides the mere speed of the folding reaction, non-cooperative temperature independent unfolding and probe-dependent kinetics have been judged as important criteria for downhill folding. For example, Gruebele and co-workers reported non-exponential kinetics under conditions with native bias in the folding of two proteins, a double-mutant of human ubiquitin and yeast phosphoglycerate kinase, and the results were interpreted as downhill folding (Sabelko et al. 1999). Furthermore, the same group suggested that the probe-dependent kinetics under a strong bias towards the native state of an engineered variant of the lambda repressor was a signature of downhill folding (Ma and Gruebele 2005). However, the relation between odd kinetics and downhill energy surfaces was recently scrutinized by Hagen (2003, 2007), who concluded that kinetic signatures proposed for downhill such as non-exponential and probe-dependent kinetics folding are unreliable, and it is thus not possible to discriminate in favor of downhill folding only from relaxation kinetics.

In addition to downhill folding taking place under strong native bias, as suggested by the energy landscape theory (Bryngelson et al. 1995), Munoz and coworkers have suggested an alternative downhill folding process. If the cancellation of enthalpic and entropic contributions to the free energy during folding is almost perfect, a protein would fold without encountering any significant energy barrier (i.e. greater than $2k_BT$ to $3k_BT$) and the folding would

proceed "downhill" in free energy landscape at any condition (Oliva and Munoz 2004). This unimodal downhill folding would only be limited by a slight ruggedness of the free-energy surface. When a protein, folding in a unimodal downhill process is subjected to denaturing conditions it is expected to unfold gradually, with the unfolded and native states merging together depending on the experimental conditions (Garcia-Mira et al. 2002). The exciting prospective of capturing partially (un)folded protein species triggered a hunt for ultra-fast folding proteins potentially displaying unimodal downhill folding.

The 35-residue subdomain from the chicken villin headpiece (VHP) is the smallest known native sequence that folds into a globular structure. Aiming to reach the folding speed limit VHP was engineered so that unfavorable electrostatic interactions in native VHP was removed and the folding rate of VHP was increased from a few μs at 300 K to 0.7 μs (Kubelka et al. 2006). Despite this record breaking folding speed, a small barrier still remains, suggesting that the onset of unimodal downhill folding is faster than anticipated.

In principle, it should be possible to distinguish between cooperative and unimodal downhill folding by NMR studies, by monitoring the folding of individual residues during thermal denaturation (Ferguson et al. 2004). This approach was employed by Sadqi et al. (2006) on a protein proposed to fold in a unimodal downhill fashion, a naphthylalanine-labelled truncated version of BBL. Proton NMR data was recorded during thermal unfolding data on 158 backbone and side-chains protons. As the midpoints of the atomic unfolding curves were not superimposable, but spanned a temperature range of 60 K, it was suggested that at any given temperature, some hydrogen atoms were in a folded environment whereas others were in an unfolded environment. However, the results were later re-analyzed by two independent groups, that both concluded that the quality of the data was not sufficient to adequately discriminate between unimodal downhill folding and folding with cooperative transition (Ferguson et al. 2007; Zhou and Bai 2007).

Finally, the hallmark of unimodal downhill folding is an underlying conformational distribution. Therefore, single-molecule experiments, such as those exemplified by Eaton and coworkers (Schuler et al. 2002; Lipman et al. 2003), may provide the final proof of the existence of a protein folding in a unimodal downhill fashion (Knott and Chan 2006; Huang et al. 2007) but until conclusive evidences of downhill folding proteins are presented, folding, as we know it, remains a barrier limited process.

Acknowledgments Y.I. is supported by a grant from the Wenner-Gren Foundations (Sweden). Work partly supported by grants from the Italian Ministero dell'Istruzione dell'Università e della Ricerca (RBIN04PWNC to M.B. and 2005027330_005 to C.T.A.) and from Sapienza Università di Roma (C26A06AFEK to M.B.).

References

Abkevich VI, Gutin AM, Shakhnovich EI (1994) Specific nucleus as the transition state for protein folding: evidence from the lattice model. Biochemistry 33:10026–10036

Anfinsen CB, Haver E, Sela M, White FHJ (1961) The kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain. Proc Natl Acad Sci USA 47:1309–1314

Bashford D, Cohen FE, Karplus M, Kuntz ID, Weaver DL (1988) Diffusion–collision model for the folding kinetics of myoglobin. Proteins 4:211–227

Bieri O, Wirz J, Hellrung B, Schutkowski M, Drewello M, Kiefhaber T (1999) The speed limit for protein folding measured by triplet–triplet energy transfer. Proc Natl Acad Sci USA 96:9597–95601

Borgia A, Bonivento D, Travaglini-Allocatelli C, Di Matteo A, Brunori M (2006) Unveiling a hidden folding intermediate in C-type cytochromes by protein engineering. J Biol Chem 281:9331–9336

Borreguero JM, Ding F, Buldyrev SV, Stanley HE, Dokholyan NV (2004) Multiple folding pathways of the SH3 domain. Biophys J 87:521–533

Brockwell DJ, Radford SE (2007) Intermediates: ubiquitous species on folding energy landscapes? Curr Opin Struct Biol 17:30–37

Bryngelson JD, Onuchic JN, Socci ND, Wolynes PG (1995) Funnels, pathways, and the energy landscape of protein folding: a synthesis. Proteins 21:167–195

Buscaglia M, Kubelka J, Eaton WA, Hofrichter J (2005) Determination of ultrafast protein folding rates from loop formation dynamics. J Mol Biol 347:657–664

Caflisch A (2004) Protein folding: simple models for a complex process. Structure 12:1750–1752

Capaldi AP, Kleanthous C, Radford SE (2002) Im7 folding mechanism: misfolding on a path to the native state. Nat Struct Biol 9:209–216

Daggett V, Fersht AR (2003) Is there a unifying mechanism for protein folding? Trends Biochem Sci 28:18–25

Eyring H (1935) The activated complex in chemical reactions. J Chem Phys 3:107–115

Ferguson N, Capaldi AP, James R, Kleanthous C, Radford SE (1999) Rapid folding with and without populated intermediates in the homologous four-helix proteins Im7 and Im9. J Mol Biol 286:1597–1608

Ferguson N, Schartau PJ, Sharpe TD, Sato S, Fersht AR (2004) One-state downhill versus conventional protein folding. J Mol Biol 344:295–301

Ferguson N, Sharpe TD, Johnson CM, Schartau PJ, Fersht AR (2007) Structural biology: analysis of 'downhill' protein folding. Nature 445:E14–E15

Fersht AR (1995) Optimization of rates of protein folding: the nucleation–condensation mechanism and its implications. Proc Natl Acad Sci USA 21:10869–10873

Fersht AR (2000) Transition-state structure as a unifying basis in protein-folding mechanisms: contact order, chain topology, stability, and the extended nucleus mechanism. Proc Natl Acad Sci USA 97:1525–1529

Fersht AR, Sato S (2004) Phi-value analysis and the nature of protein-folding transition states. Proc Natl Acad Sci USA 101:7976–7981

Fersht AR, Matouschek A, Serrano L (1992) The folding of an enzyme. I. Theory of protein engineering analysis of stability and pathway of protein folding. J Mol Biol 224:771–782

Friel CT, Capaldi AP, Radford SE (2003) Structural analysis of the rate-limiting transition states in the folding of Im7 and Im9:

similarities and differences in the folding of homologous proteins. J Mol Biol 326:293–305
Friel CT, Beddard GS, Radford SE (2004) Switching two-state to three-state kinetics in the helical protein Im9 via the optimisation of stabilising non-native interactions by design. J Mol Biol 342:261–273
Garcia-Mira MM, Sadqi M, Fischer N, Sanchez-Ruiz JM, Muñoz V (2002) Experimental identification of downhill protein folding. Science 298:2191–2195
Gianni S, Travaglini-Allocatelli C, Cutruzzolà F, Bigotti MG, Brunori M (2001) Snapshots of protein folding. A study on the multiple transition state pathway of cytochrome c(551) from *Pseudomonas aeruginosa*. J Mol Biol 309:1177–1187
Gianni S, Guydosh NR, Khan F, Caldas TD, Mayor U, White GW, DeMarco ML, Daggett V, Fersht AR (2003a) Unifying features in protein-folding mechanisms. Proc Natl Acad Sci USA 100:13286–13291
Gianni S, Travaglini-Allocatelli C, Cutruzzolà F, Brunori M, Shastry MC, Roder H (2003b) Parallel pathways in cytochrome c(551) folding. J Mol Biol 330:1145–1152
Gianni S, Calosci N, Aelen JM, Vuister GW, Brunori M, Travaglini-Allocatelli C (2005) Kinetic folding mechanism of PDZ2 from PTP-BL. Protein Eng Des Sel 18:389–395
Gianni S, Brunori M, Travaglini-Allocatelli C (2007a) Plasticity of the protein folding landscape: switching between on- and off-pathway intermediates. Arch Biochem Biophys 466:172–176
Gianni S, Geierhaas CD, Calosci N, Jemth P, Vuister GW, Travaglini-Allocatelli C, Vendruscolo M, Brunori M (2007b) A PDZ domain recapitulates a unifying mechanism for protein folding. Proc Natl Acad Sci USA 104:128–133
Gianni S, Ivarsson Y, Jemth P, Brunori M, Travaglini-Allocatelli C (2007c) Identification and characterization of protein folding intermediates. Biophys Chem 128:105–113
Gorski SA, Capaldi AP, Kleanthous C, Radford SE (2001) Acidic conditions stabilise intermediates populated during the folding of Im7 and Im9. J Mol Biol 312:849–863
Hagen SJ (2003) Exponential decay kinetics in "downhill" protein folding. Proteins 50:1–4
Hagen SJ (2007) Probe-dependent and nonexponential relaxation kinetics: unreliable signatures of downhill protein folding. Proteins 68:205–217
Huang CY, Klemke JW, Getahun Z, DeGrado WF, Gai F (2001) Temperature-dependent helix-coil transition of an alanine based peptide. J Am Chem Soc 123:9235–9238
Huang F, Sato S, Sharpe TD, Ying L, Fersht AR (2007) Distinguishing between cooperative and unimodal downhill protein folding. Proc Natl Acad Sci USA 104:123–127
Hubner IA, Lindberg M, Haglund E, Oliveberg M, Shakhnovich EI (2006) Common motifs and topological effects in the protein folding transition state. J Mol Biol 359:1075–1085
Itoh K, Sasai M (2006) Flexibly varying folding mechanism of a nearly symmetrical protein: B domain of protein A. Proc Natl Acad Sci USA 103:7298–7303
Itzhaki LS, Otzen DE, Fersht AR (1995) The structure of the transition state for folding of chymotrypsin inhibitor 2 analysed by protein engineering methods: evidence for a nucleation–condensation mechanism for protein folding. J Mol Biol 254:260–288
Ivarsson Y, Travaglini-Allocatelli C, Jemth P, Malatesta F, Brunori M, Gianni S (2007) An on-pathway intermediate in the folding of a PDZ domain. J Biol Chem 282:8568–8572
Jackson SE (1998) How do small single-domain proteins fold? Fold Des 3:R81–91
Jackson SE, Fersht AR (1991) Folding of chymotrypsin inhibitor 2. 1. Evidence for a two-state transition. Biochemistry 30:10428–10435
Jemth P, Gianni S, Day R, Li B, Johnson CM, Daggett V, Fersht AR (2004) Demonstration of a low-energy on-pathway intermediate in a fast-folding protein by kinetics, protein engineering, and simulation. Proc Natl Acad Sci USA 101:6450–6455
Juraszek J, Bolhuis PG (2006) Sampling the multiple folding mechanisms of Trp-cage in explicit solvent. Proc Natl Acad Sci USA 103:15859–15864
Karplus M, Weaver DL (1976) Protein-folding dynamics. Nature 260:404–406
Karplus M, Weaver DL (1994) Protein folding dynamics: the diffusion–collision model and experimental data. Protein Sci 3:650–668
Kiefhaber T (1995) Kinetic traps in lysozyme folding. Proc Natl Acad Sci USA 92:9029–9033
Knott M, Chan HS (2006) Criteria for downhill protein folding: calorimetry, chevron plot, kinetic relaxation, and single-molecule radius of gyration in chain models with subdued degrees of cooperativity. Proteins 65:373–391
Krantz BA, Mayne L, Rumbley J, Englander SW, Sosnick TR (2002) Fast and slow intermediate accumulation and the initial barrier mechanism in protein folding. J Mol Biol 324:359–371
Kubelka J, Hofrichter J, Eaton WA (2004) The protein folding 'speed limit'. Curr Opin Struct Biol 14:76–88
Kubelka J, Chiu TK, Davies DR, Eaton WA, Hofrichter J (2006) Sub-microsecond protein folding. J Mol Biol 359:546–553
Lam AR, Borreguero JM, Ding F, Dokholyan NV, Buldyrev SV, Stanley HE, Shakhnovich E (2007) Parallel folding pathways in the SH3 domain protein. J Mol Biol 373:1348–1360
Lapidus LJ, Eaton WA, Hofrichter J (2002) Measuring dynamic flexibility of the coil state of a helix-forming peptide. J Mol Biol 319:19–25
Lazaridis T, Karplus M (1997) "New view" of protein folding reconciled with the old through multiple unfolding simulations. Science 278:1928–1931
Levinthal C (1968) Are there pathways for protein folding? J Chem Phys 65:44–45
Lindberg MO, Oliveberg M (2007) Malleability of protein folding pathways: a simple reason for complex behaviour. Curr Opin Struct Biol 17:21–29
Lindberg M, Tangrot J, Oliveberg M (2002) Complete change of the protein folding transition state upon circular permutation. Nat Struct Biol 9:818–822
Lipman EA, Schuler B, Bakajin O, Eaton WA (2003) Single-molecule measurement of protein folding kinetics. Science 301:1233–1235
Ma H, Gruebele M (2005) Kinetics are probe-dependent during downhill folding of an engineered lambda6-85 protein. Proc Natl Acad Sci USA 102:2283–2287
Matouschek A, Kellis JT Jr, Serrano L, Bycroft M, Fersht AR (1990) Transient folding intermediates characterized by protein engineering. Nature 346:440–445
Mayor U, Grossmann JG, Foster NW, Freund SM, Fersht AR (2003a) The denatured state of Engrailed homeodomain under denaturing and native conditions. J Mol Biol 333:977–991
Mayor U, Guydosh NR, Johnson CM, Grossmann JG, Sato S, Jas GS, Freund SMV, Alonso DOV, Daggett V, Fersht AR (2003b) The complete folding pathway of a protein from nanoseconds to microseconds. Nature 421:863–867
McCallister EL, Alm E, Baker D (2000) Critical role of beta-hairpin formation in protein G folding. Nat Struct Biol 7:669–673
Munoz V, Thompson PA, Hofrichter J, Eaton WA (1997) Folding dynamics and mechanism of beta-hairpin formation. Nature 390:196–199
Myers JK, Oas TG (1999) Contribution of a buried hydrogen bond to lambda repressor folding kinetics. Biochemistry 38:6761–6768
Neira JL, Davis B, Ladurner AG, Buckle AM, Gay Gde P, Fersht AR (1996) Towards the complete structural characterization of a protein folding pathway: the structures of the denatured, transition and native states for the association/folding of two complementary

fragments of cleaved chymotrypsin inhibitor 2. Direct evidence for a nucleation–condensation mechanism. Fold Des 1:189–208

Nishimura C, Prytulla S, Dyson JH, Wright PE (2000) Conservation of folding pathways in evolutionarily distant globin sequences. Nat Struct Biol 7:679–686

Oliva FY, Munoz V (2004) A simple thermodynamic test to discriminate between two-state and downhill folding. J Am Chem Soc 126:8596–8597

Onuchic JN, Socci ND, Luthey-Schulten Z, Wolynes PG (1996) Protein folding funnels: the nature of the transition state ensemble. Fold Des 1:441–450

Onuchic JN, Luthey-Schulten Z, Wolynes PG (1997) Theory of protein folding: the energy landscape perspective. Annu Rev Phys Chem 48:545–600

Otzen DE, Fersht AR (1997) Folding of circular and permuted chymotrypsin inhibitor 2: retention of the folding nucleus. J Mol Biol 37:8139–8146

Rao F, Settanni G, Guarnera E, Caflisch A (2005) Estimation of protein folding probability from equilibrium simulations. J Chem Phys 122:184901

Religa TL, Markson JS, Mayor U, Freund SM, Fersht AR (2005) Solution structure of a protein denatured state and folding intermediate. Nature 437:1053–1056

Religa TL, Johnson CM, Vu DM, Brewer SH, Dyer RB, Fersht AR (2007) The helix-turn-helix motif as an ultrafast independently folding domain: the pathway of folding of engrailed homeodomain. Proc Natl Acad Sci USA 104:9272–9277

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

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

Sanchez IE, Kiefhaber T (2003) Evidence for sequential barriers and obligatory intermediates in apparent two-state protein folding. J Mol Biol 325:367–376

Sato S, Fersht AR (2007) Searching for multiple folding pathways of a nearly symmetrical protein: temperature dependent phi-value analysis of the B domain of protein a. J Mol Biol 372:254–267

Schuler B, Lipman EA, Eaton WA (2002) Probing the free-energy surface for protein folding with single-molecule fluorescence spectroscopy. Nature 419:743–747

Spudich GM, Miller EJ, Marqusee S (2004) Destabilization of the *Escherichia coli* RNase H kinetic intermediate: switching between a two-state and three-state folding mechanism. J Mol Biol 335:609–618

Travaglini-Allocatelli C, Gianni S, Brunori M (2004) A common folding mechanism in the cytochrome c family. Trends Biochem Sci 29:535–541

Wallin S, Zeldovich KB, Shakhnovich EI (2007) The folding mechanics of a knotted protein. J Mol Biol 368:884–893

Weikl TR, Dill KA (2007) Transition-states in protein folding kinetics: the structural interpretation of Phi values. J Mol Biol 365:1578–1586

White GW, Gianni S, Grossmann JG, Jemth P, Fersht AR, Daggett V (2005) Simulation and experiment conspire to reveal cryptic intermediates and a slide from the nucleation–condensation to framework mechanism of folding. J Mol Biol 350:757–775

Williams S, Causgrove TP, Gilmanshin R, Fang KS, Callender RH, Woodruff WH, Dyer RB (1996) Fast events in protein folding: helix melting and formation in a small peptide. Biochemistry 35:691–697

Wolynes PG (2004) Latest folding game results: protein A barely frustrates computationalists. Proc Natl Acad Sci USA 101:6837–6838

Wong KB, Clarke J, Bond CJ, Neira JL, Freund SM, Fersht AR, Daggett V (2000) Towards a complete description of the structural and dynamic properties of the denatured state of barnase and the role of residual structure in folding. J Mol Biol 296:1257–1282

Wright CF, Lindorff-Larsen K, Randles LG, Clarke J (2003) Parallel protein-unfolding pathways revealed and mapped. Nat Struct Biol 10:658–662

Zhou Z, Bai Y (2007) Structural biology: analysis of protein-folding cooperativity. Nature 445:E16–E17