Toso *et al. Biology Direct* (2025) 20:44
https://doi.org/10.1186/s13062-025-00642-x

Biology Direct

RESEARCH Open Access

Topological determinants in protein folding dynamics: a comparative analysis of metamorphic proteins

Julian Toso[1†], Valeria Pennacchietti[1†], Mariana Di Felice[1], Eduarda S. Ventura[1], Angelo Toto[1] and Stefano Gianni[1*]

Abstract
Protein folding remains a fundamental challenge in molecular biology, particularly in understanding how polypeptide chains transition from denatured states to their functional conformations. Here we analyze the folding mechanisms of the engineered metamorphic proteins B4 and Sb3, which share highly similar sequences but adopt distinct topologies. Kinetic analyses revealed that B4 follows a two-state folding mechanism, whereas Sb3 involves the formation of an intermediate species. We further explore the role of topology in folding commitment using the metamorphic mutant Sb4, which can populate both conformations. By analyzing folding and unfolding behaviors under varying experimental conditions, our findings suggest that topology dictates folding mechanisms at an early stage. These results demonstrate that folding landscapes are primarily shaped by final native structures rather than sequence composition.

Keywords Protein folding, Kinetics, Metamorphic proteins, Site-directed mutagenesis

Introduction

The classical description of the protein folding problem invoked two apparently inter-related challenges: the prediction of a protein structure from its amino acid sequence and the detailed characterization of the mechanism whereby polypeptide chains achieve their native, functionally competent, conformations. Whilst the recent developments of AlphaFold artificial intelligence tools have contributed to remarkable breakthroughs in protein structure predictions [1], addressing the sequence of events which lead the denatured conformation into its native counterpart remains a major problem in molecular biology [2–6]. The significance of shedding light onto such processes is further reinforced by noting that the characterization of folding mechanisms holds the potential to uncover fundamental insights into the links between protein folding and misfolding, a recurrent and devastating event in several human diseases.

The Holy Grail of biophysical studies would be to derive general principles through the analysis of simple systems. This quest is particularly challenging in the context of protein folding, as a critical complication lies in comparing the data obtained from proteins displaying different structure and sequence. Indeed, the complexity arises not only from variations in protein sequences among different proteins but also from the structural diversity observed within both native and denatured states. A common strategy to circumvent this problem lies in comparing the folding mechanisms of homologous proteins, i.e. proteins sharing a similar structure whilst displaying

†Julian Toso and Valeria Pennacchietti contributed equally to this work.

*Correspondence:
Stefano Gianni
stefano.gianni@uniroma1.it
[1]Dipartimento di Scienze Biochimiche "A. Rossi Fanelli", Sapienza Università di Roma, Laboratory affiliated to Istituto Pasteur Italia - Fondazione Cenci Bolognetti, P.le Aldo Moro 5, Rome 00185, Italy

different sequences [7–13]. These kinds of studies, which have been extensively discussed elsewhere, have contributed to building an invaluable framework which offers significant insights into the interplay between sequence and structure in protein folding, enhancing our understanding of this complex process.

A complementary perspective to the traditional comparative analysis of homologous proteins is provided by the study of metamorphic proteins [14–17]. These proteins retain the highest possible sequence identity while exhibiting distinct changes in their protein topology, thereby allowing to expand the current understanding of structural versatility and functional adaptation in protein biology. Nevertheless, experimental works on the folding of metamorphic proteins are, to date, very limited.

Bryan, Orban and coworkers recently reported a highly ambitious experiment of protein design that successfully evolved different pairs of metamorphic proteins in vitro [18]. Such proteins were obtained by engineering the original sequence of the ribosomal protein S6 from *Thermus thermophilus* (101 amino acids in length) into converging high-identity sequences, which could populate three distinct folds: the very same fold of the ribosomal protein S6 - referred to as S-fold - or the folds adopted by two domains of streptococcal Protein G, the B-fold of the IgG-binding GB domain (56 amino acids in length), and the A-fold of the HSA-binding GA domain (56 amino acids in length). Among the engineered variants, the B4 and Sb3 variants stand out as particularly intriguing for protein folding studies, as these two constructs display nearly identical sequences between residues 1 and 56, differing only at position 5, where B4 has a leucine, while Sb3 has a tyrosine, and present an overlapping αββ topology between residues 23 and 56 (Fig. 1A). Despite these similarities, the presence of 31 additional amino acids on the C-terminal end of Sb3 leads to a markedly different structure compared to B4. Specifically, while B4 assumes the B-fold found in the GB domain of Protein G (Fig. 1B), Sb3 adopts the characteristic S-fold of the ribosomal protein S6 (Fig. 1C). This remarkable protein engineering achievement offers the relatively unique opportunity to perform a complementary study on protein folding mechanisms, allowing for two key inter-related questions to be directly addressed: first, can any common intermediates be encountered along the folding pathway of proteins displaying long, highly identical stretches of amino acids? Secondly, at which stage does a protein commit to its native state?

Here we provide a comparative analysis of the mechanism of folding of B4 and Sb3. We show that, despite nearly identical sequences from position 1 to 56 and a partially overlapping secondary structure topology, these two constructs fold via different mechanisms, corresponding to a two-state scenario and a three-state scenario for B4 and Sb3, respectively. Analysis of folding and unfolding rate constants under varying experimental conditions indicates that the intermediate observed in Sb3 exhibits distinct stability characteristics compared to B4. This suggests that their folding pathways diverge early in the process and that the B4 fold does not resemble the intermediate species observed for Sb3. To further investigate this divergence, we also characterized the folding of the Y5L mutant of Sb3 - referred to as Sb4 for simplicity - which was previously shown to populate simultaneously both the B4 and Sb3 conformations. Notably, the folding behavior of Sb4 mirrors the two distinct phases observed individually in B4 and Sb3, with each phase responding to experimental conditions in a manner consistent with its corresponding construct. As discussed below, these findings provide a strong validation of the notion that protein topology is committed very early along the folding pathway [19, 20] and, thus, that folding mechanisms are primarily dictated by topology rather than by amino acid sequence.

Results

The kinetics of folding of B4 and Sb3 reveal the presence of divergent folding mechanisms

The primary objective of the present study was to deliver a comprehensive comparative analysis of the folding mechanism of the B4 and Sb3 engineered proteins. Consequently, both constructs were subjected to stopped-flow (un)folding kinetics, by rapidly diluting the protein into solutions containing different guanidine hydrochloride (GdnHCl) concentrations. In all experiments, the time course of the emission fluorescence was satisfactorily fitted to a single-exponential equation. The logarithm of the observed rate constants k_{obs} plotted against denaturant concentration, commonly referred to as the chevron plot, obtained for B4 and Sb3 are reported in Fig. 1D and E, respectively. In both cases, the chevron conformed to a simple V-shape, consistent with a two-state scenario [21].

It is of interest to compare the folding behavior of B4 and Sb3 with those of their parent proteins, GB and S6 [22–24]. In fact, in both cases, the folding reaction seems marginally affected by the pronounced changes in aminoacidic composition, as evidenced by the nearly identical observed folding rate constant as well as by a very similar dependence on denaturant concentration, namely m_F value (Fig. 1D and E). However, the engineered constructs exhibited dramatically higher unfolding rate constants, with an increase of 60-fold for B4 (k_u = 35.9 ± 1.4 s^{-1}) compared to GB1 (k_u = 0.57 ± 0.02 s^{-1}) and of 40,000-fold for Sb3 (k_u = 12.4 ± 1.2 s^{-1}) compared to S6 (k_u = 0.0003 ± 0.00004 s^{-1}). Such observation indicates a pronounced destabilization in both cases (Fig. 1D and E).

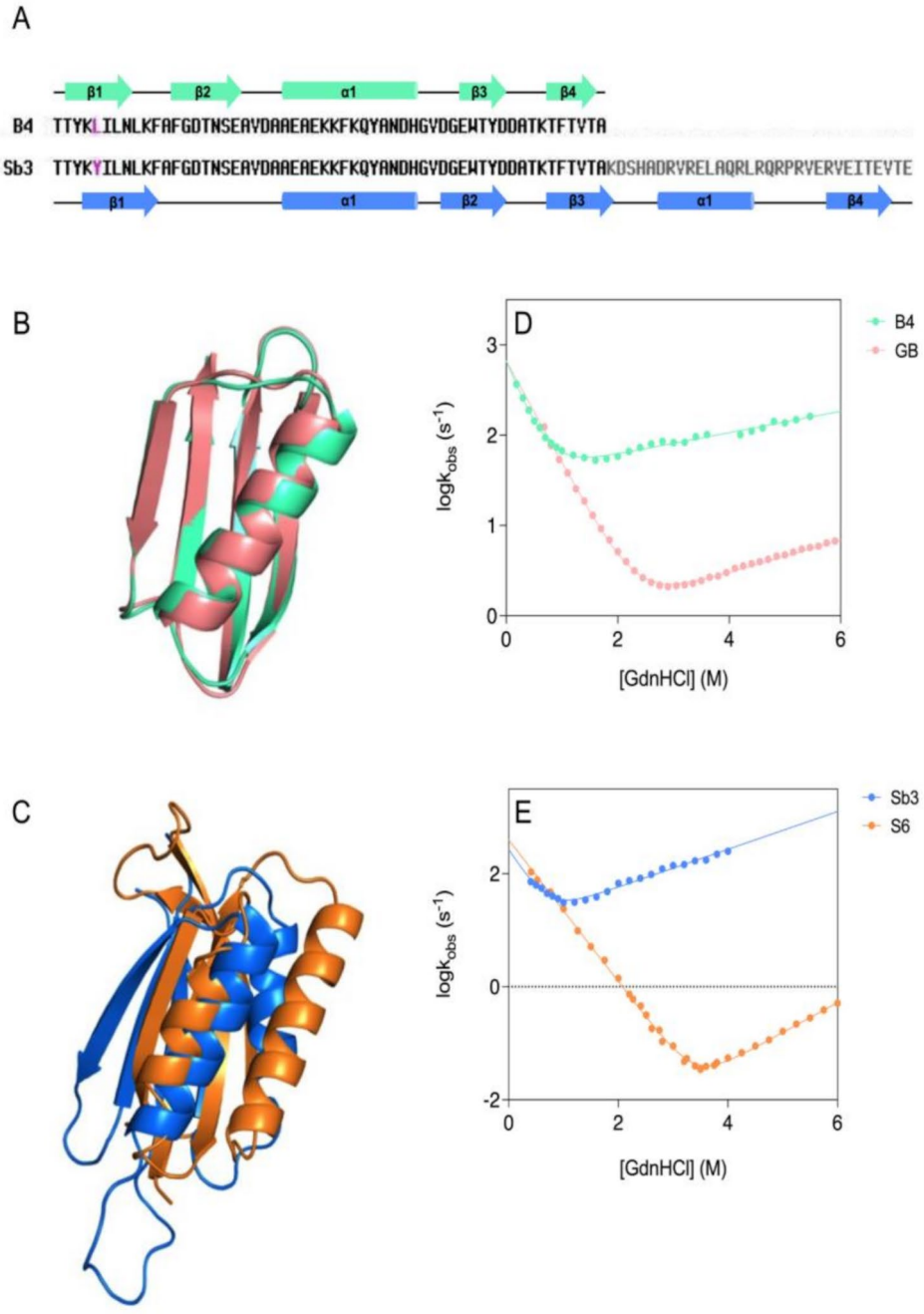

Fig. 1 Comparative analysis of (un)folding kinetics for B4, Sb3, and their parent proteins GB and S6. (**A**) Sequence and topology alignment of the B4 and Sb3 constructs, differing by a single residue within the first 56 amino acids (adapted from Ruan et al., 2023). (**B**) Structural alignment of the B4 construct (green) with the parent IgG-binding GB domain (pink) from Streptococcal Protein G (PDB: 1PGA). (**C**) Structural alignment of the Sb3 construct (blue) with its parent ribosomal protein S6 (orange) from *T. thermophilus* (PDB: 1FKA). (**D**) Comparison of the chevron plots obtained for the B4 construct (green) and the GB domain (pink). Both exhibit classical V-shaped two-state folding behavior, with similar kinetic parameters except for a significantly higher unfolding rate constant in B4. (**E**) Comparison of the chevron plots obtained for the Sb3 construct (blue) and the S6 protein (orange). Both display classical V-shaped two-state folding behavior, with similar kinetic parameters except for a significantly higher unfolding rate constant in Sb3. The (un)folding kinetic data for the GB domain and the S6 protein were taken from Morrone et al., 2011, and Otzen and Oliveberg, 1999, respectively

In an effort to overcome the marginal stability of the engineered constructs, folding and unfolding experiments were also performed in the presence of 0.3 M Na_2SO_4, a commonly used salt to stabilize protein. Remarkably, despite successfully improving the stability of both proteins, the addition of 0.3 M Na_2SO_4 produced markedly different effects on the B4 and Sb3 constructs. Specifically, while the chevron plot of B4 retained its

characteristic V-shaped profile (Fig. 2A), indicative of a two-state folding mechanism, the chevron plot of Sb3 revealed a clear deviation from linearity in its refolding arm (i.e., at low denaturant concentrations), conventionally denoted as roll-over effect [25, 26]. This suggests the existence of a three-state folding scenario, determined by the accumulation of an intermediate species in rapid pre-equilibrium with the unfolded state, which was not detectable in the absence of Na_2SO_4 (Fig. 2B). We note that these findings parallel what had been previously observed on the S6 protein [24], further confirming that the engineered constructs strictly mirror the folding pathways of their respective wild-type counterparts.

Investigation of the (un)folding kinetics and thermodynamic properties of B4 and Sb3 under varying experimental conditions reveals early divergence in their respective folding pathways

A powerful strategy for uncovering the structural and stability characteristics of rapid pre-equilibrium intermediates and transition state ensembles (TSE) involved in folding mechanisms lies in measuring the folding and unfolding reactions, while changing the experimental conditions. To this end, time-resolved fluorescence monitored (un)folding experiments were performed at different pH conditions for both B4 and Sb3, in the presence of stabilizing salt (Fig. 3). Under all investigated conditions, the chevron plots of B4 were accurately fitted with a two-state model. On the other hand, the chevron plots of Sb3 required a kinetic three-state model to account for the persistent "roll-over effect" observed at all pHs. Table 1 summarizes the kinetic and thermodynamic parameters obtained from the global fits of B4 and Sb3 across all experimental conditions.

Inspection of Fig. 3 reveals that, whilst the refolding arm of B4 shows minimal sensitivity to changes in pH, its unfolding is significantly affected outside of the pH range 4.5-7.0 (Fig. 3A). This results in a pH dependence of the (un)folding equilibrium rate constant displaying two distinct sigmoidal, monotonic transitions at both acidic and alkaline pH values (Fig. 3B). This behavior is consistent with the protonation of at least two groups in the native state, with apparent pKa values of approximately 4 and 8. It is particularly noteworthy to note that the GB domain of Protein G, the parent protein of B4, was previously shown to display a very similar dependence of protein stability with changing pH [22].

Contrary to the behavior observed for B4, the chevron plots of Sb3 showed minimal sensitivity to changes in pH, with the profiles remaining largely unaltered (Fig. 3C). It is evident that both the unfolding and refolding arms do not display significant fluctuations. Figure 3D highlights the negligible change of the stability of the folding intermediate as a function of pH. On the basis of this observation, we can unambiguously conclude that the intermediate species of Sb3 does not resemble the B4 native fold.

The study of the (un)folding of the metamorphic variant Sb4 helps in addressing the role of sequence and/or topology in the early commitment of a protein folding pathway

As briefly recalled above, B4 and Sb3 display identical sequence composition between position 1 and 56, with the sole exception of one residue at position 5 (Fig. 1).

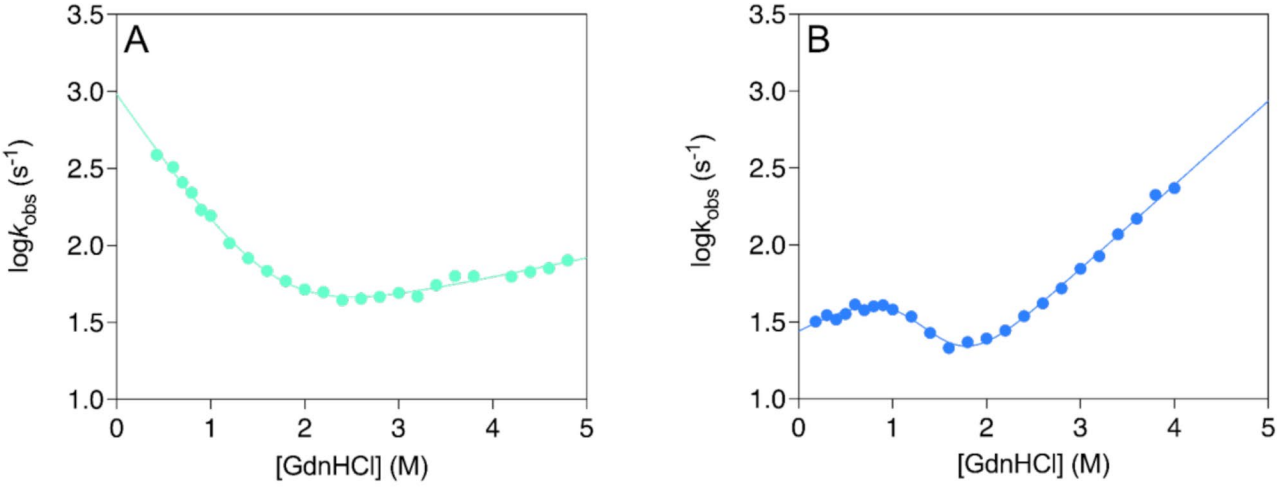

Fig. 2 Supplementation of 0.3 M Na_2SO_4 to the 50 mM Hepes pH 7.5 experimental Buffer highlights the distinct folding pathways pursued by the B4 and Sb3 Constructs. (**A**) Chevron plot of the (un)folding kinetics of the B4 construct in 50 mM Hepes pH 7.5 buffer supplemented with 0.3 M Na_2SO_4, showing a two-state folding scenario. (**B**) Chevron plot of the (un)folding kinetics of Sb3 in 50 mM Hepes pH 7.5 buffer supplemented with 0.3 M Na_2SO_4 reveals a three-state folding mechanism, characterized by the appearance of a previously undetected intermediate species in rapid pre-equilibrium with the unfolded state

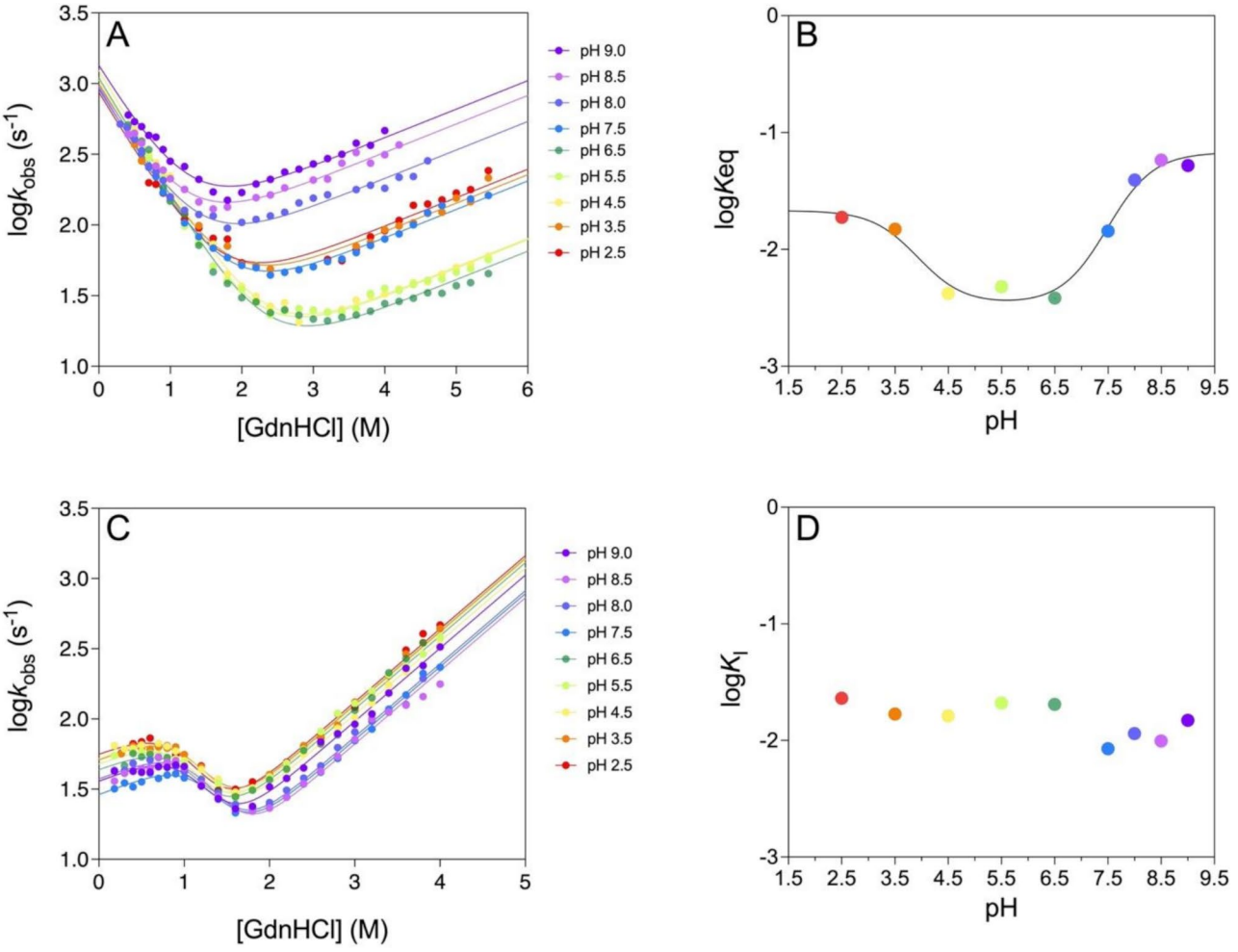

Fig. 3 Comparative analysis on the pH-dependent folding behavior of B4 and Sb3. (**A**) Chevron plots of B4 measured from pH 2.5 to pH 9.0. Lines are the best global fit to a two-state equation with shared m-values. (**B**) Logarithm of the unfolding equilibrium constants between the native and denatured state of B4 versus pH. The fit clearly reveals the presence of two monotonic sigmoidal transitions over pH 7 and below pH 4.5, respectively. (**C**) Chevron plots of Sb3 measured from pH 2.5 to pH 9.0. Lines are the best global fit to a three-state equation with shared m-values. (**D**) Logarithm of the partition factor K_I of Sb3, proportional to the difference between the activation barriers for the pre-equilibrium intermediate state to revert to the denatured state rather than proceeding to the native state, versus pH

Table 1 Kinetic parameters of the folding of B4 and Sb3 measured at different pH conditions

pH	B4			Sb3		
	k_f (s^{-1})	k_u (s^{-1})	K_{Eq} (10^{-2})	k_f (s^{-1})	k_u (s^{-1})	K_I (10^{-2})
9.0	1270 ± 62	66 ± 3	5.0 ± 0.3	34 ± 2	2.4 ± 0.1	1.5 ± 0.1
8.5	900 ± 48	52 ± 2	6.0 ± 0.3	34 ± 2	1.7 ± 0.1	1.0 ± 0.2
8.0	871 ± 37	34 ± 2	4.0 ± 0.2	35 ± 2	1.9 ± 0.1	1.0 ± 0.1
7.5	910 ± 43	13.0 ± 0.8	1.0 ± 0.1	27 ± 1	1.8 ± 0.1	0.80 ± 0.05
6.5	1085 ± 45	4.0 ± 0.2	0.40 ± 0.02	41 ± 2	3.0 ± 0.1	2.0 ± 0.1
5.5	1048 ± 43	5 ± 0.3	0.50 ± 0.02	45 ± 3	3.1 ± 0.2	2.0 ± 0.1
4.5	1221 ± 50	5 ± 0.3	0.40 ± 0.02	50 ± 2	2.7 ± 0.1	1.6 ± 0.1
3.5	975 ± 43	14.0 ± 0.8	1.00 ± 0.07	48 ± 3	3.2 ± 0.2	1.7 ± 0.2
2.5	844 ± 41	16.0 ± 0.9	2.00 ± 0.01	53 ± 3	3.4 ± 0.2	2.0 ± 0.1

As reported above the folding parameters obtained for B4 and Sb3 were calculated from a two-state and three-state model, respectively. All data were recorded at 25 °C. To reduce errors, the chevron plots of each protein were globally fitted with shared m-values

Mutation of tyrosine 5 to leucine in Sb3, explicitly referred to as the Sb4 variant, results in a partial shift in the native population from the S-fold towards the B-fold [18]. In fact, the HSQC NMR spectrum for this variant revealed roughly twice the number of amide cross-peaks compared to B4 and Sb3, while closely matching the amide signals produced by each protein construct in their respective NMR spectra. This finding strongly indicates that the Sb4 variant is a metamorphic protein capable of populating both the B- and S-folds simultaneously [18].

The unique structural features of Sb4 provide an invaluable opportunity to explore the impact of native state topology on protein folding pathways, in the context of the very same protein. Hence, we resorted to investigate the kinetic (un)folding reactions of Sb4 under varying pH conditions and in the presence of stabilizing salt, following the same approach that we applied to the B4 and Sb3 variants.

Contrary to what was observed in the case of both B4 and Sb3, when subjecting the Sb4 protein to time-resolved folding experiments, we could clearly observe double-exponential behavior (Fig. 4, Supplementary Fig. 1). Such distinct phases could be detected both in folding and unfolding experiments. The dependence of the fast and slow kinetic phases is reported in Fig. 5. Of interest, it is evident that the overall behavior of Sb4 captures both the chevron plots of B4 and Sb3. Along with previously reported NMR data, this finding is a signature of the coexistence of two slow-interconverting conformational populations: (i) the Sb4_B-fold, exhibiting the characteristic two-state profile of B4; (ii) the Sb4_S-fold, displaying the three-state mechanism of Sb3. Indeed, comparison of the chevron plots obtained from each of the three engineered constructs under different pH conditions revealed that the folding behavior of Sb4_B-fold and Sb4_S-fold is affected in the same way as that of B4 and Sb3, respectively. Table 2 summarizes the kinetic and thermodynamic parameters obtained from the global fits of Sb4_B-fold and Sb4_S-fold across all experimental conditions.

Specifically, as with the B4 construct, the unfolding arms of Sb4_B-fold are strongly influenced by alkaline pH values above 7.5 (Fig. 5A). On the other hand, the acidic conditions that impact B4 have no effect on Sb4_B-fold, resulting in the logarithmic plots of equilibrium constants over pH showing one single, monotonic transition at alkaline pH values exceeding 7.5 (Fig. 5B). In contrast, in Sb4_S-fold, the tested pH conditions have no significant effect on the kinetic constants of (un)folding and the intermediate (Fig. 5C and D), just like observed in Sb3.

Discussion

The classical studies on protein folding are generally limited by the selection of a specific protein system, which inherently limits the possibility to extract general rules. A common strategy to overcome these complications lies in comparing the folding of (i) protein homologues, sharing the same topology while displaying a different sequence [7–10, 12, 13, 27], or (ii) proteins with different structure but with the highest possible sequence identity [17, 20]. In this context, the peculiar sequence identity between

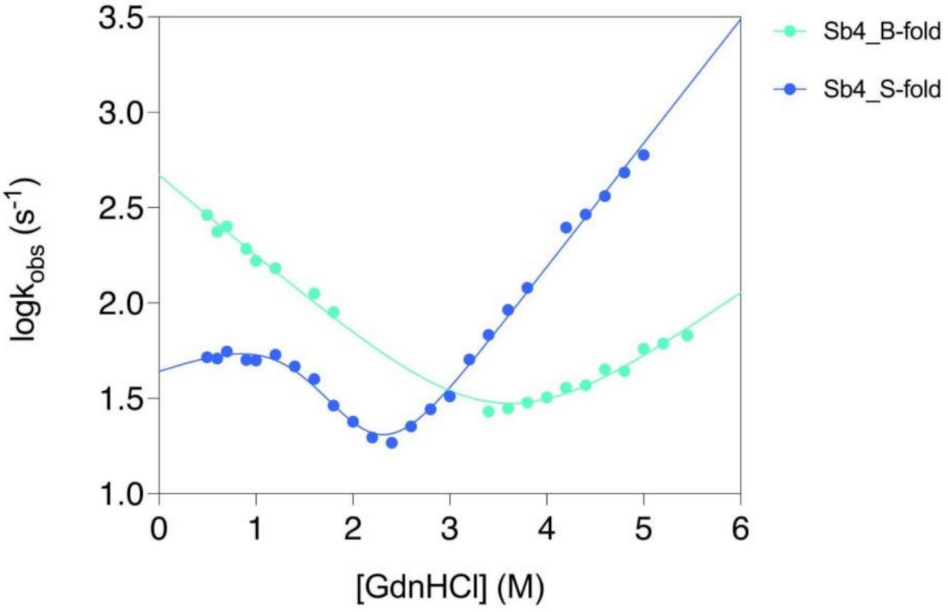

Fig. 4 Chevron plots of the two distinct phases observed for the Sb4 variant in 50 mM Bis-Tris pH 6.5 buffer supplemented with 0.3 M Na_2SO_4. The two phases correspond to the Sb4_B-fold and Sb4_S-fold populations, whose profiles strongly mirror the behavior of B4 and Sb3, respectively

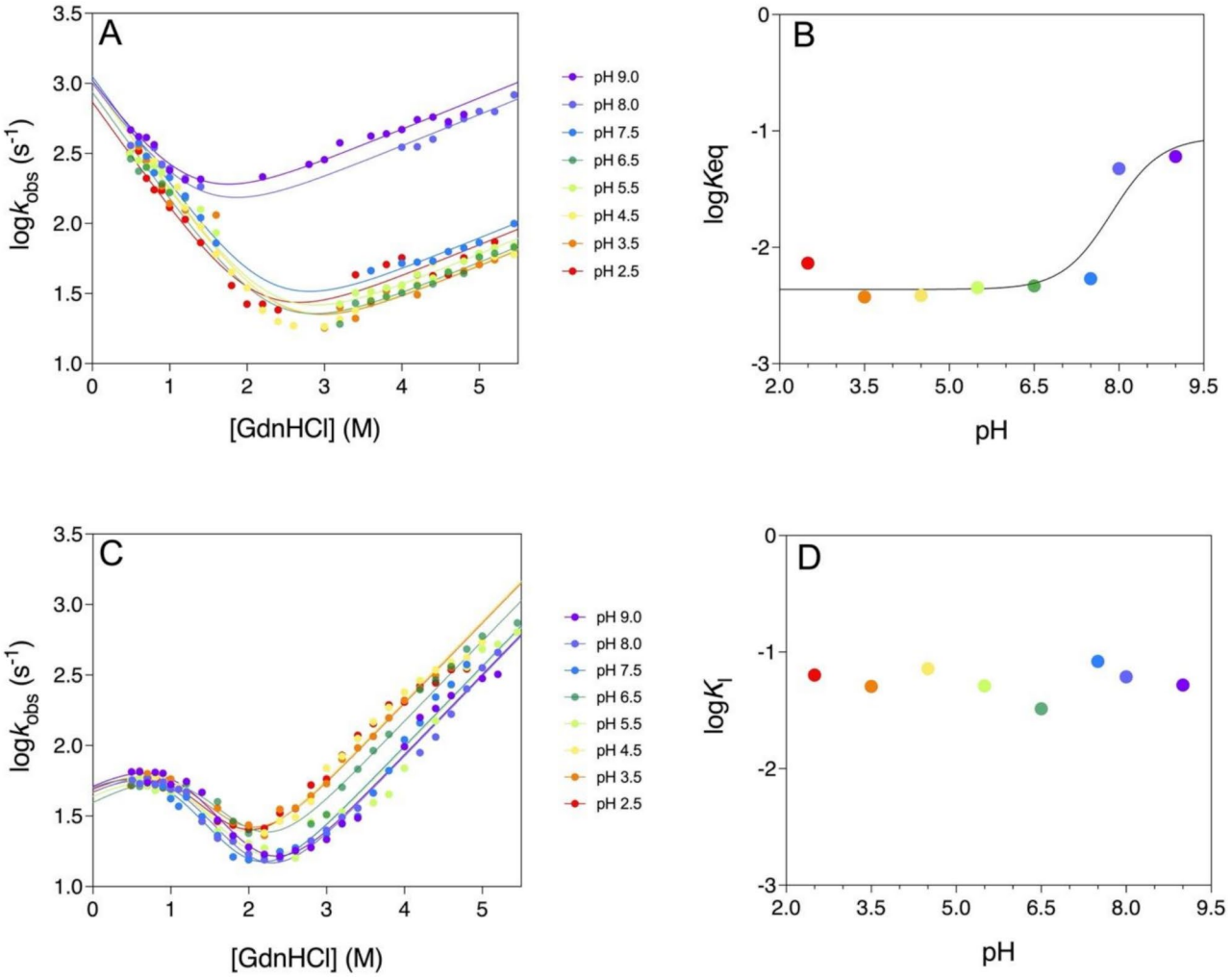

Fig. 5 Analysis of the pH-dependent folding behavior of the Sb4 variant. (**A**) Chevron plots of the Sb4_B-fold population measured from pH 2.5 to pH 9.0. Lines are the best global fit to a two-state equation with shared m-values. (**B**) Logarithm of the unfolding equilibrium constants between the native and denatured state of Sb4_B-fold versus pH. The fit clearly reveals the presence of a single, monotonic sigmoidal transition over pH 7.5 (**C**) Chevron plots of the Sb4_S-fold population measured from pH 2.5 to pH 9.0. Lines are the best global fit to a three-state equation with shared m-values. (**D**) Logarithm of the partition factor KI of Sb4_S-fold, proportional to the difference between the activation barriers for the pre-equilibrium intermediate state to revert to the denatured state rather than proceeding to the native state, versus pH

Table 2 Kinetic parameters of the two distinct phases of Sb4 variant, representing the B-fold and S-fold, measured at different pH condition

pH	Sb4 _ B-fold			Sb4 _ S-fold		
	k_f (s^{-1})	k_u (s^{-1})	K_{Eq} (10^{-2})	k_f (s^{-1})	k_u (s^{-1})	K_I (10^{-2})
9.0	994 ± 95	60 ± 7	6.0 ± 0.6	53 ± 13	0.45 ± 0.06	5.0 ± 0.3
8.0	962 ± 100	46 ± 6	5.0 ± 0.6	49 ± 12	0.44 ± 0.05	6.0 ± 0.4
7.5	1116 ± 94	6.0 ± 0.8	0.50 ± 0.05	54 ± 15	0.52 ± 0.06	8.0 ± 0.3
6.5	863 ± 70	4.0 ± 0.5	0.50 ± 0.04	40 ± 11	0.79 ± 0.10	3.00 ± 0.06
5.5	1024 ± 81	4.6 ± 0.6	0.4 ± 0.04	45 ± 11	0.5 ± 0.06	5.0 ± 0.2
4.5	1004 ± 87	3.9 ± 0.5	0.4 ± 0.04	53 ± 14	1.08 ± 0.10	7.0 ± 0.4
3.5	1012 ± 86	3.8 ± 0.5	0.4 ± 0.04	48 ± 12	1.05 ± 0.10	5.0 ± 0.1
2.5	733 ± 60	5.4 ± 0.7	0.7 ± 0.07	51 ± 15	1.08 ± 0.10	6.0 ± 0.4

As reported above the folding parameters obtained for Sb4_B-fold and S-fold were calculated from a two-state and three-state model, respectively. All data were recorded at 25 °C. To reduce errors, the chevron plots of each condition were globally fitted with shared m-values

B4 and Sb3 (Fig. 1) offers the unique opportunity to provide a complementary experimental perspective.

The experimental characterization reported above demonstrates that, whilst B4 folds via a robust two-state kinetic mechanism, stabilization of Sb3 allows for the detection of a meta-stable folding intermediate. Such behavior parallels what observed earlier for their wild-type counterparts, namely GB1 for B4 and S6 for Sb3 [22–24]. Given that the first 56 amino acids of Sb3 are essentially identical to the sequence of B4,with the sole exception of one residue at position 5, a question of interest is whether the folding intermediate of Sb3 displays similar structural features as those of B4. This insight is valuable as it may reveal the point at which folding pathways diverge and become committed to the native topology, thereby offering a glimpse of the width of the free energy landscape at the early stages of folding. In this perspective, it is particularly instructive to discuss the changes on stability as probed by different pH conditions. In fact, whilst the folding of B4 is highly sensitive to pH, the intermediate of Sb3 appears extremely robust, with nearly identical chevron plots from pH 2.5 to 9.0. This finding allows to conclude unequivocally that the intermediate of Sb3 has distinct structural properties as compared to B4 and, therefore, that the folding pathways of such proteins diverge very early. We note that this finding supports the notion that folding pathways commit very early to native topologies, which might be already imprinted as early as in the denatured state [20].

One of the most fascinating theories to describe the folding of proteins implies the native state to be minimally frustrated [28–30]. A consequence of the principle of minimal frustration is that the energy landscape is typically funneled towards the native state and alternative structures are highly unstable. Furthermore, in the case of minimally frustrated sequences, a small set of collective parameters that measure proximity to the native state can effectively indicate the energy of a given configuration [31–33]. Consequently, the final structure of a protein is expected to play a crucial role in shaping the topology of the folding landscape, as frequently observed in the comparison of homologous proteins. The analysis of the folding pathway of Sb4 variant provides an exquisite validation of this prediction. In fact, since this variant assumes two alternative native conformations in a slow equilibrium, it offers the opportunity to test the robustness of the links between folding pathway and native topology on the very same protein. A comparison between the folding behavior of Sb4, B4 and Sb3 highlights an unambiguous robustness in the folding pathways leading to the S- and B-folds. Hence, alternative topologies demand specific folding routes, even in the presence of the same amino acidic sequence.

Conclusions

The comparative analysis of B4, Sb3, and Sb4 proteins provides compelling evidence that protein topology plays a decisive role in dictating folding mechanisms. Despite nearly identical sequences, the divergent folding pathways of B4 and Sb3, along with the dual folding behavior of Sb4, demonstrate that the commitment to a specific native state occurs early in the folding process. The observed kinetic differences and stability variations suggest that structural constraints imposed by the final topology significantly influence the folding landscape. By offering a unique perspective on metamorphic proteins, this study emphasizes the importance of native-state topology in guiding folding pathways.

Materials and methods

Plasmid construction

The cDNA sequences encoding for the Sb4 protein [18] was subcloned between the NdeI and BamHI restriction sites of pET28b(+) plasmid, a vector bearing a kanamycin resistance marker which is commonly used for the expression of N-terminally 6xHis-tagged constructs. The eXact tag pH0720 expression vectors encoding the B4 and Sb3 proteins [18] were a kind gift from Prof. Philip N. Bryan. This vector closely resembles the pPal8 expression vector (Biorad), bearing an ampicillin resistance marker, with the only key difference being that the original Profinity eXact tag found in the pPal18 vector is replaced in pH0720 by the optimized N-terminal fusion subtilisin prodomain sequence:

MEAVDANSLA QAKEAAIKEL KQYGIGDKYI KLINNAKTVE GVESLKNEIL KALPTEGSGN TIRVIVSVDKAKFNPHEVLG IGGHIVYQFK LIPAVVVDVP ANAVGKLKKM PGVEKVEFDH QYRGL which gives much better soluble expression compared to the original version [18]. The transformations for DNA extraction were performed in E. Coli XL10-Gold Ultracompetent Cells (Agilent) and all constructs were verified by DNA sequencing (Eurofins Genomics).

Protein expression and purification

Plasmids were transformed in *E. Coli* BL21(DE3) Competent Cells (Thermo Scientific) by heat shock procedure. Bacterial cells containing the recombinant pET28b(+) or eXact tag pH0720 vectors were grown in Miller LB Broth (Sigma Aldrich) supplemented with either 30 μg/ml of kanamycin or 100 μg/ml of ampicillin, respectively, at 37 °C and 180 rpm. When bacterial cultures reached an OD_{600} of 0.6–0.8, protein expression was induced by adding 1 mM isopropyl β-d-1-thio-galactopyranoside (IPTG) and was allowed to proceed overnight at 25 °C and 180 rpm. Cells were harvested by centrifugation for 10 min at 7000 g and 4 °C. Bacterial pellets deriving from expression of the Sb4 protein encoded in the pET28b(+)

vector were resuspended in 50 mM TrisHCl buffer at pH 7.5, 300 mM NaCl, 10 mM imidazole (Buffer A), whereas bacterial pellets deriving from expression of the eXact tag pH0720 vector were resuspended in 100 mM sodium phosphate buffer at pH 7.2 (Buffer P), supplemented in both cases with cOmplete™ EDTA-free Protease Inhibitor Cocktail (Roche). Cell lysis was performed by sonication and cellular debris were removed by ultracentrifugation for 45 min at 18,600 g and 6 °C. The soluble fraction of bacterial lysate containing Sb4 protein was loaded onto a 5 ml nickel charged HiTrap chelating HP column (Cytiva) equilibrated with Buffer A. The column was washed with 10 column volumes of 50 mM TrisHCl buffer at pH 7.5, 1 M NaCl to remove any unspecific binding and proteins were successively eluted by performing a 0 M to 1 M imidazole gradient through an ÄKTA-start (Cytiva) HPLC system. Fractions containing the protein of interest were collected and the buffer was exchanged to 50 mM TrisHCl at pH 7.5, 300 mM NaCl, using a HiTrap Desalting column (Cytiva) according to the manufacturer's instructions. The soluble fraction of bacterial lysate containing either prodomain fusion proteins B4 or Sb3 was loaded on a 5 ml EconoFit Profinity eXact column (Bio-Rad) equilibrated with Buffer P. The column was washed with 10 column volumes of Buffer *P* to remove any unspecific binding and was successively equilibrated with 2 column volumes of 100 mM sodium phosphate buffer at pH 7.2, 100 mM NaF, (Buffer E) and incubated for 30 min at 25 °C to allow for the catalytic activation of the immobilized protease within the matrix. The cleaved target proteins were eluted using Buffer E. The purity of all protein samples was verified through SDS-PAGE.

Kinetic (un)folding experiments

Rapid mixing (un)folding experiments were performed using a SX-18 single mixing stopped-flow device (Applied Photophysics). All experiments were performed at 25 °C, with a final protein concentration of 1 μM and GdnHCl as the denaturing agent. Fluorescence emissions were measured with a 360 nm cut-off glass filter, using an excitation wavelength of 280 nm. When possible, at least five individual traces were acquired and then averaged for each experiment. Buffers used for salt-dependence of B4 and Sb3 were 50 mM Hepes pH 7.5, supplemented with either 0 mM or 300 mM Na_2SO_4. Buffers used for pH-dependence of B4, Sb3 and Sb4 were 50 mM Phosphate pH 2.5, 50mM Formate pH 3.5, 50mM Acetate pH 4.5–5.5, 50 mM Bis-Tris pH 6.5, 50mM Hepes pH 7.0, 50mM TrisHCl pH 8.0-8.5-9.0 all supplemented with 300 mM Na_2SO_4. All chevron plots of the B4 protein were fitted with a standard two-state equation:

$$\log k_{obs} = \log\left(k_u e^{(m_u[GdnHCl]/RT)} + k_f e^{(-m_f[GdnHCl]/RT)}\right)$$

Chevron plots of the Sb3 protein were fitted to a three-state equation implying the accumulation of an intermediate in rapid pre-equilibrium with the denatured state:

$$\log k_{obs} = \log\left(k_u e^{(m_u[GdnHCl]/RT)} + \frac{k_f e^{(-m_f[GdnHCl]/RT)}}{1 + K_I e^{(m_I[GdnHCl]/RT)}}\right)$$

with the exception of the experiment performed in 50 mM Hepes pH 7.5, without salt, which was fitted to a standard two-state equation. The chevron plots of the two distinct phases which could be recorded in the experiments involving the Sb4 protein were fitted either with a two-state or three-state equation, depending on the presence of a clear roll-over effect in the refolding arm. By following a standard procedure in protein folding studies, to minimize statistical errors all the chevron plots were analyzed with shared kinetic *m*-values.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13062-025-00642-x.

Supplementary Material 1

Author contributions

J.T., V.P., M.d.F., E.V., performed research; J.T., V.P., A.T. and S.G., conceived research; J.T. V.P. and S.G. analyzed the data; A.T. and S.G. supervised research; J.T., V.P. and S.G. wrote the first version of the manuscript and all authors read and approved the final version.

Funding

This work was partly supported by grants from European Union's MSCA Doctoral Networks under the Grant Agreement IDPro No 101,119,633 (to S.G.), by grants from Sapienza University of Rome (RG12017297FA7223, RG1231888A88129D to S.G., RG124190E53A81B8, RM12218148DA1933 to A.T.), the Associazione Italiana per la Ricerca sul Cancro (Individual Grant—IG 24551 to S.G.), the Istituto Pasteur Italia ("Research Program 2022 to 2023 Under 45 Call 2020" to A.T.), the Italian MUR-PRIN 2022 grant N. 2022JY3PMB to A.T. We acknowledge co-funding from Next Generation EU, in the context of the National Recovery and Resilience Plan, and the Investment PE8—Project Age-It: "Ageing Well in an Ageing Society". This resource was co-financed by the Next Generation EU [DM 1557 11 October 2022]. The views and opinions expressed are only those of the authors and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the European Commission can be held responsible for them.

Data availability

No datasets were generated or analysed during the current study.

Declarations

Competing interests

The authors declare no competing interests.

Received: 26 February 2025 / Accepted: 24 March 2025
Published online: 03 April 2025

References

1. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A, Bridgland A, Meyer C, Kohl SAA, Ballard AJ, Cowie A, Romera-Paredes B, Nikolov S, Jain R, Adler J, Back T, Petersen

S, Reiman D, Clancy E, Zielinski M, Steinegger M, Pacholska M, Berghammer T, Bodenstein S, Silver D, Vinyals O, Senior AW, Kavukcuoglu K, Kohli P, Hassabis D. Highly accurate protein structure prediction with alphafold. Nature. 2021;596:583–9.
2. Bhatia S, Udgaonkar JB. Understanding the heterogeneity intrinsic to protein folding. Curr Opin Struct Biol. 2024;84:102738.
3. Fersht AR. From covalent transition States in chemistry to noncovalent in biology: from β- to Φ-value analysis of protein folding. Q Rev Biophys. 2024;57:e4.
4. Gianni S, Brunori M. The folding and misfolding of multidomain proteins. Mol Asp Med. 2024;101:101337.
5. Jemth P. Protein binding and folding through an evolutionary lens. Curr Opin Struct Biol. 2025;90:102980.
6. Kannan A, Naganathan AN. Engineering the native ensemble to tune protein function: diverse mutational strategies and interlinked molecular mechanisms. Curr Opin Struct Biol. 2024;89:102940.
7. Calosci N, Chi CN, Richter B, Camilloni C, Engstrom Å, Eklund L, Travaglini-Allocatelli C, Gianni S, Vendruscolo M, Jemth P. Comparison of successive transition States for folding reveals alternative early folding pathways of two homologous proteins. Proc Natl Acad Sci U S A. 2008;105:p19241–19246.
8. Clarke J, Cota E, Fowler SB, Hamill SJ. Folding studies of Ig-like beta-sandwich proteins suggest they share a common folding pathway. Structure. 1999;7:1145–53.
9. Ferguson N, Capaldi AP, James R, Kleanthous C, Radford SE. Rapid folding with and without populated intermediates in the homologous four-helix proteins Im7 and Im9. J Mol Biol. 1999;286:1597–608.
10. Gianni S, Guydosh NR, Khan F, Caldas TD, Mayor U, White GW, DeMarco ML, Daggett V, Fersht AR. Unifying features in protein-folding mechanisms. Proc Natl Acad Sci USA. 2003;100(23):13286–91.
11. Nandi T, Yadav A, Ainavarapu SRK. Experimental comparison of energy landscape features of ubiquitin family proteins. Proteins. 2020;88:449–61.
12. Scott KA, Randles LG, Clarke J. The folding of spectrin domains II: phi-value analysis of R16. J Mol Biol. 2004;344:207–11.
13. Travaglini-Allocatelli C, Gianni S, Morea V, Tramontano A, Soulimane T, Brunori M. Exploring the cytochrome C folding mechanism: cytochrome c552 from thermus thermophilus folds through an on-pathway intermediate. J Biol Chem. 2003;278(42):41136–40.
14. Alexander A PA, He Y, Chen Y, Orban J, Bryan N PN. The design and characterization of two proteins with 88% sequence identity but different structure and function. Proc Natl Acad Sci U S A. 2007;104:11963–8.
15. Alexander A PA, He Y, Chen Y, Orban J, Bryan N PN. A minimal sequence code for switching protein structure and function. Proc Natl Acad Sci U S A. 2009;106:21149–54.
16. Bryan PN, Orban J. Implications of protein fold switching. Curr Opin Struct Biol. 2013;23:314–6.
17. Giri R, Morrone A, Travaglini-Allocatelli C, Jemth P, Brunori M, Gianni S. Folding pathways of proteins with increasing degree of sequence identities but different structure and function. Proc Natl Acad Sci U S A. 2012;109:p17772–6.
18. Ruan B, He Y, Chen Y, Choi EJ, Chen Y, Motabar D, Solomon T, Simmerman R, Kauffman T, Gallagher DT, Orban J, Bryan PN. Design and characterization of a protein fold switching network. Nat Commun. 2023;14:431.
19. Lindorff-Larsen K, Røgen P, Paci E, Vendruscolo M, Dobson CM. Protein folding and the organization of the protein topology universe. Trends Biochem Sci. 2005;30:13–9.
20. Morrone A, McCully ME, Bryan PN, Brunori M, Daggett V, Gianni S, Travaglini-Allocatelli C. The denatured state dictates the topology of two proteins with almost identical sequence but different native structure and function. J Biol Chem. 2011;286:3863–72.
21. Itzhaki LS, Otzen DE, Fersht AR. 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. 1995;254(2):260–88.
22. Morrone A, Giri R, Toofanny RD, Travaglini-Allocatelli C, Brunori M, Daggett V, Gianni S. GB1 is not a two-state folder: identification and characterization of an on-pathway intermediate. Biophys J. 2011;101:2053–60.
23. Otzen DE, Kristensen O, Proctor M, Oliveberg M. Structural changes in the transition state of protein folding: alternative interpretations of curved chevron plots. Biochemistry. 1999;38:6499–511.
24. Otzen DE, Kristensen O, Proctor M. Structural changes in the transition state of protein folding: alternative interpretations of curved chevron plots. Biochemistry. 1999;38:6499–511.
25. Capaldi AP, Shastry MC, Kleanthous C, Roder H, Radford SE. Ultrarapid mixing experiments reveal that Im7 folds via an on-pathway intermediate. Nat Struct Biol. 2001;8(1):68–72.
26. Parker MJ, Spencer J, Clarke AR. An integrated kinetic analysis of intermediates and transition States in protein folding reactions. J Mol Biol. 1995;253(5):771–86.
27. Zarrine-Afsar A, Larson SM, Davidson AR. The family feud: do proteins with similar structures fold via the same pathway? Curr Opin Struct Biol. 2005;15(1):42–9.
28. Bryngelson JD, Onuchic JN, Socci ND, Wolynes PG. Funnels, pathways, and the energy landscape of protein folding: a synthesis. Proteins. 1995;21:167–95.
29. Onuchic JN, Socci ND, Luthey-Schulten Z, Wolynes PG. Protein folding funnels: the nature of the transition state ensemble. Fold Des. 1996;1:441–50.
30. Wolynes PG. Energy landscapes and solved protein-folding problems. Philos Transact Roy Soc Math Phys Eng Sci. 2005;363:453–64.
31. Ferreiro DU, Hegler JA, Komives EA, Wolynes PG. Localizing frustration in native proteins and protein assemblies. Proc Natl Acad Sci U S A. 2007;104:p19819–19824.
32. Ferreiro DU, Hegler JA, Komives EA, Wolynes PG. On the role of frustration in the energy landscapes of allosteric proteins. Proc Natl Acad Sci USA. 2011;108:p3499–3503.
33. Oliveberg M, Wolynes PG. The experimental survey of protein-folding energy landscapes. Q Rev Biophys. 2005;38:245–88.

Publisher's note