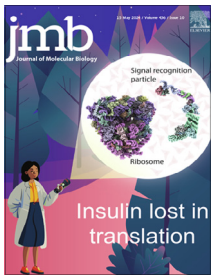

The Mechanism of Folding of Human Frataxin in Comparison to the Yeast Homologue – Broad Energy Barriers and the General Properties of the Transition State

Paola Pietrangeli [1,†], Lucia Marcocci [1,†], Valeria Pennacchietti [1], Awa Diop [1], Mariana Di Felice [1], Livia Pagano [1], Francesca Malagrinò [2], Angelo Toto [1], Maurizio Brunori [1] and Stefano Gianni [1,*]

*1 - **Dipartimento di Scienze Biochimiche** "**A. Rossi Fanelli**", Sapienza Università di Roma, Laboratory Affiliated to Istituto Pasteur Italia - Fondazione Cenci Bolognetti, 00185 Rome, Italy*

*2 - **Dipartimento di Medicina Clinica,** Sanità Pubblica, Scienze Della Vita e Dell'ambiente, Università dell'Aquila, Piazzale Salvatore Tommasi 1, 67010 L'Aquila – Coppito, Italy*

***Correspondence to Stefano Gianni:** stefano.gianni@uniroma1.it (S. Gianni)*

https://doi.org/10.1016/j.jmb.2024.168555

Edited by Monika Fuxreiter

Abstract

The funneled energy landscape theory suggests that the folding pathway of homologous proteins should converge at the late stages of folding. In this respect, proteins displaying a broad energy landscape for folding are particularly instructive, allowing inferring both the early, intermediate and late stages of folding. In this paper we explore the folding mechanisms of human frataxin, an essential mitochondrial protein linked to the neurodegenerative disorder Friedreich's ataxia. Building upon previous studies on the yeast homologue, the folding pathway of human frataxin is thoroughly examined, revealing a mechanism implying the presence of a broad energy barrier, reminiscent of the yeast counterpart. Through an extensive site-directed mutagenesis, we employed a Φ -value analysis to map native-like contacts in the folding transition state. The presence of a broad energy barrier facilitated the exploration of such contacts in both early and late folding events. We compared results from yeast and human frataxin providing insights into the impact of native topology on the folding mechanism and elucidating the properties of the underlying free energy landscape. The findings are discussed in the context of the funneled energy landscape theory of protein folding.

Introduction

The classical description of reaction mechanisms demands the detailed representation of the entire process, leading from the reactants to the products. This aim would imply the characterization of the composition, structure and energy of all the intermediates and intervening transition states.[1] In the case of protein folding, however, the extraordinary co-operative nature of the reaction makes this task particularly difficult. Indeed, whilst hundreds of weak non-covalent bonds between the chain and the solvent are formed and broken, the experimental manifestation of folding is often highly co-operative and only the fully native and denatured states might be detected.

A powerful approach to enrich our knowledge on the protein folding process is to compare the folding of proteins sharing a similar topology while displaying a different amino acid composition. In fact, the comparative analysis on these systems may help in drawing general rules on the nature of transition states, as well as in establishing the robustness of folding pathways with changing

amino acid sequences. Consequently, different experimental projects have focussed on the description of the folding of different members of the same protein family.[2–7]

Frataxin is an essential mitochondrial protein involved in the metabolism of iron and responsible for the human neurodegenerative disease Friedrech's ataxia.[8–12] The folding pathway of yeast frataxin has been previously extensively studied and it was shown that this protein folds via a complex mechanism, which may be associated to the presence of a broad free energy barrier between the native and denatured state.[13,14] This feature is particularly interesting since it allows to elude, at least in part, the experimental limitations arising from the co-operative nature of folding and permits to infer a characterization of both the early and late events of folding.[15,16]

Here we provide a comprehensive characterization of the folding pathway of human frataxin and compare it with previous results obtained on the yeast homologue. By taking advantage of extensive site directed mutagenesis we show that also human frataxin folds via a broad energy barrier as discovered for the yeast homologue,[13,14] despite only the latter displays a detectable cold denaturation at temperatures higher than 0 °C.[17] Then, by performing a so-called Φ-value analysis, we map the presence of native like structural clusters in the folding transition state. The fortuitous presence of a broad energy barrier allowed us to successfully address the structural features of both the late and early events of folding. As detailed below, the obtained results offer the rare opportunity to scrutinize the effects of native topology in biasing the mechanism of folding and provide a glimpse of the properties of the underlying energy landscape. Our data are therefore discussed on the light of the so-called funnelled energy landscape theory of protein folding.[18,19]

Results

To address the mechanism of folding of human frataxin we performed a Φ value analysis[20,21] and compared the results obtained to those previously published on the yeast homologue.[14] Thirty-two site-directed mutants were produced and subjected to kinetic folding experiments. Mutants were designed following the standard rules of Φ -value analysis.[21]

Human frataxin folds via a broad energy barrier

Kinetic folding experiments were performed by 11-fold dilution of native protein into a urea buffer at different concentration of denaturant (unfolding) or by rapidly mixing denatured into urea buffer solutions at different concentration of denaturant (refolding). At all the investigated conditions, experimental traces were satisfactorily fitted to single-exponential time courses at all final denaturant concentrations. The semilogarithmic plots of the observed unfolding and refolding rate constants versus denaturant concentration (chevron plots) for human frataxin and its mutants are shown in Figure 1. It is evident that, in analogy to what previously observed in the yeast variant, the measured chevron plots are characterized by a detectable curvature both in the folding and unfolding arms.

Various models have been classically invoked to describe the presence of deviations from linearity in chevron plots.[15,22–24] In general, non-linearity in either the unfolding or refolding limb of these plots was attributed to the existence of intermediates within the reaction pathway.[6,23–25] This complexity could either result from the rapid accumulation of such intermediates or from changes in the rate-limiting step as denaturant concentration varied, with both scenarios leading to identical mathematical solutions. Alternatively, it has been suggested that a broad transition state exists between the native and denatured states,[15,16,26,27] It is important to note that, whilst these models are often nearly experimentally indistinguishable, a kinetic test has been introduced to distinguish between these scenarios.[28] Specifically, by analyzing how the observed curvature changes under different reaction conditions, it becomes possible to reconstruct the shifts in the transition state as a function of protein stability across a broad range of stabilities. This analysis allows for the identification of more sensitive indicators, or "fingerprints," that can differentiate between various barrier profiles.

In practice, the observed chevron plots obtained at different conditions, or after site-directed mutagenesis, may be fitted to the following quadratic equation:

$$k_{obs} = k_F \cdot e^{(m_F \cdot [denaturant] + m' [denaturant]^2)} + k_U \cdot e^{(m_U \cdot [denaturant] + m' [denaturant]^2)}$$

In this equation, k_F and k_U denote the folding and unfolding rate constants in water, m_F and m_U are the respective slopes of the folding and unfolding limbs, and the degree of curvature in the chevron plot is quantified by the parameter *m*'. A perfectly V-shaped, two-state chevron plot would have *m*' close to zero. By assessing the dependence of m' under changing conditions (e.g., mutagenesis), it is possible to draw conclusions about whether the rate-limiting barrier follows a smooth, broad energy profile. In fact, whilst the three-state model would imply the presence of a bell shaped dependence of *m*' (see for example panel e in Figure 3 of ref[28]), in the case of broad energy barriers, *m*' is expected to be maintained around an average value.

Figure 2 reports the measured *m*' for human frataxin and its site-directed variants as a function of the observed thermodynamic stability. It is evident that the observed *m*' parameter does not

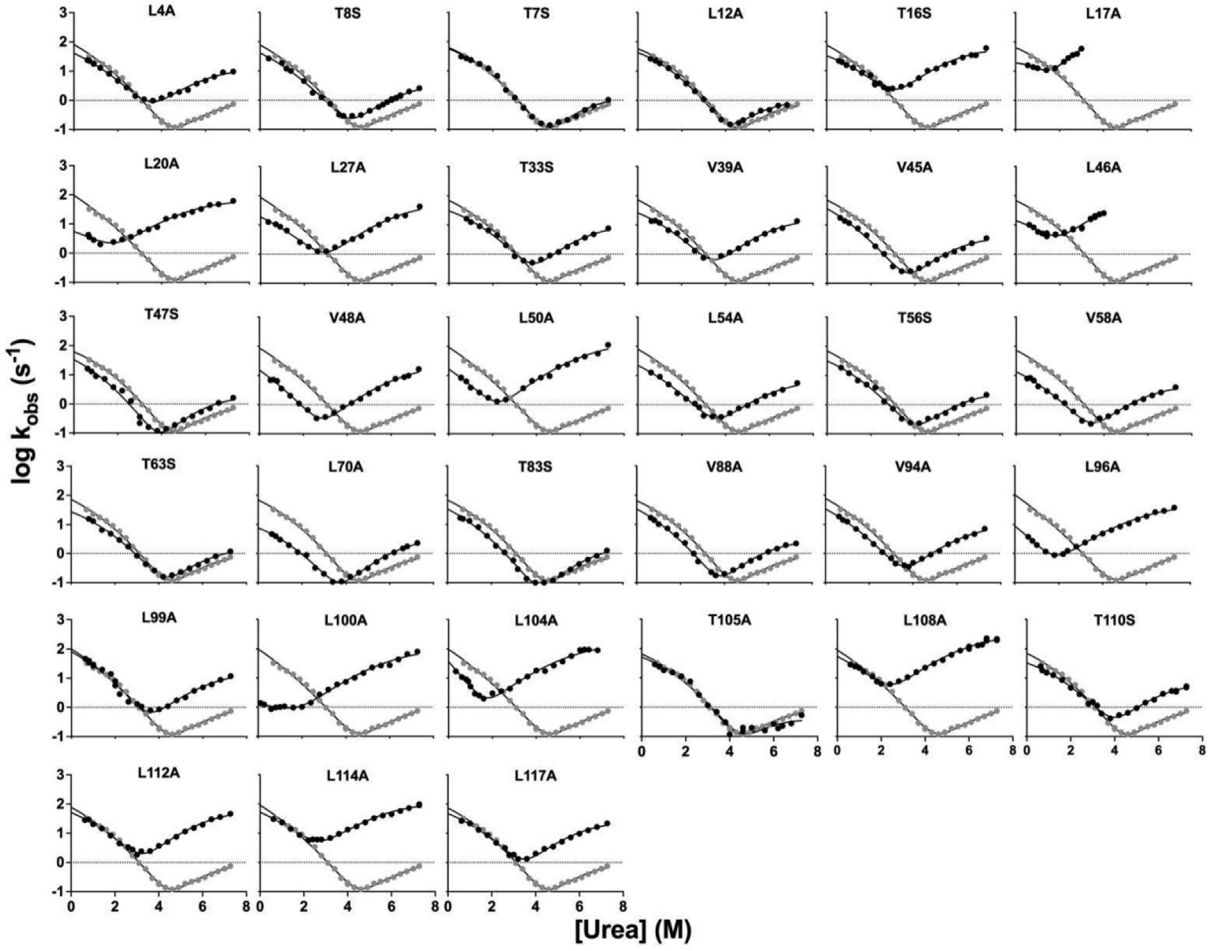

Figure 1. Chevron plots of site-directed variants of wt (in gray) and site-directed variants (in black) of human frataxin. Lines represent the best curve of a quadratic equation, taking into account curvature in both unfolding and refolding arm of the chevron plots, and describing a folding mechanism characterized by the presence of a broad energy barrier (see text for details).

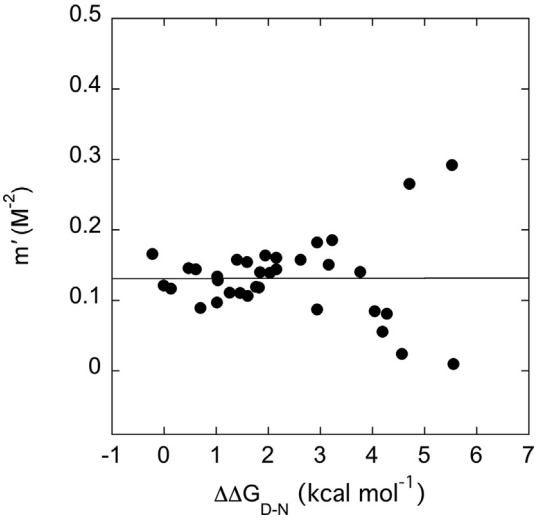

Figure 2. Calculated m' values for human frataxin site-directed variants as a function of the thermodynamic stability.

display any evident dependence on protein stability, as opposed to what expected in the presence of a folding intermediate, with an *m*' scattering around an average value of 0.14 kcal mol^{-1} M^{-2}. On the basis of these observations, in analogy to what previously proposed in the case of yeast frataxin,[14] it is possible to conclude that the folding of human frataxin most likely conforms to a broad free energy barrier.

The structural features of the folding transition states of human frataxin – early, intermediate and late events of folding

As briefly described above, the kinetics of folding of human frataxin resembles what previously observed for the yeast homologue with a broad energy barrier between the denatured and native state characterizes it. This model suggests that the complexity observed in the chevron plot is a result of progressive changes in the structure of

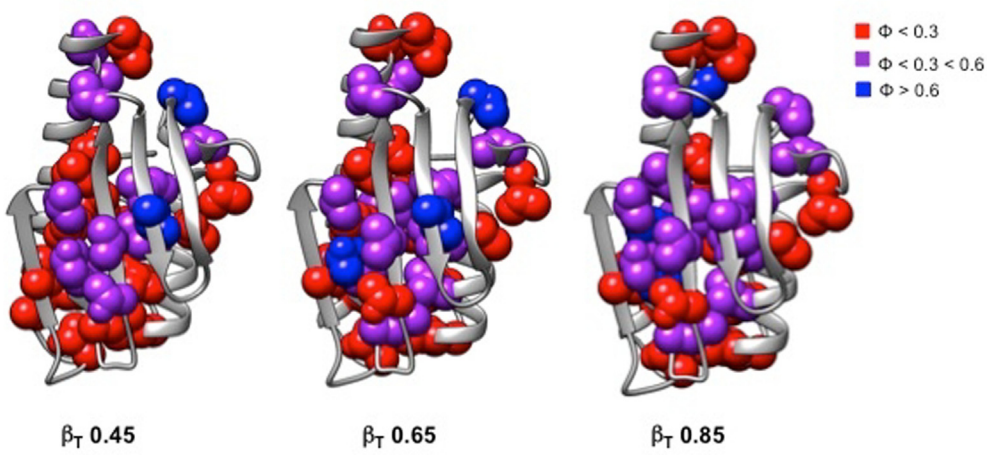

Figure 3. Color coded Φ-values calculated at different β_T are mapped on the three-dimensional structure of human frataxin (PDB: 1EKG). The analysis reveals, in the early events of folding, higher Φ-values in β-sheet of the protein, with a general increase of native-like contacts as the native state is approached. A structural superposition between human and yeast frataxin is presented on the right.

the transition state as the stability of the native state decreases.[15,16,26,27] In fact, when the transition state is associated with a broad barrier, it becomes responsive to variations in experimental conditions, such as changes in denaturant concentrations. Under these circumstances, when the protein is destabilized by the addiction of denaturant, it adopts a conformation that more closely resembles the native state, in agreement with a so-called Hammond effect.[29] In this case, the calculated Tanford's β_T value, which is an index reflecting the buried surface area of the transition state relative to the native and denatured states, changes as a function of denaturant.

To probe the structure of the transition state, we calculated reliable Φ values at various positions along the reaction coordinate, corresponding to different values of β_T. In fact, as initially championed by Oliveberg and colleagues,[15,16,26,27] the broad barrier model facilitates the exploration of the transition state across a wide range of conditions and β_T values. Importantly, it should be stressed that the Φ value analysis assumes that mutations exert no effects on the pre-exponential factor of folding. Own to the complexity of the model, it is essential to exercise caution and refrain from making extrapolations. Therefore, our experimental analysis focused on the folding of frataxin under conditions that could be directly measured, specifically at the most denatured-like and native-like conditions (i.e., at β_T values of 0.45 and 0.85). Additionally, we included the value of 0.65 to examine an intermediate scenario between these two extremes.

The Φ values for the early (β_T = 0.45), intermediate (β_T = 0.65), and late (β_T = 0.85) stages of folding, as well as the changes in free energy associated with unfolding upon mutation, are detailed in Table 1. Furthermore, a color-coded representation of the measured values mapped on the native structure of frataxin is reported in Figure 3. It is evident that the beta-sheet of the protein appears to act as a folding nucleus, with values of intermediate and high Φ sparsely distributed along this secondary structure element that may be detected in the early events of folding. Structure gradually tapers off towards the two major helices that display low values of Φ. The comparison between the values obtained for the early (β_T = 0.45), intermediate (β_T = 0.65), and late (β_T = 0.85) stages of folding highlights a consolidation of such elements of structures with a general increase of the magnitude of Φ values as the native state is approached.

A straightforward approach to compare the folding of two proteins lies in analysing the Φ-Φ correlation between homologues residues.[30] Figure 4 and Table 2 compare the Φ values obtained in the early, intermediate and late stages of folding for human and yeast frataxin. It is evident that, whilst the late stages of folding are highly robust, with the two proteins conforming to a straight line with a slope of one, scatter substantially increases when the early stages are considered (panels A and B). This finding demonstrates that there is a weak native bias at the early stages of folding, as opposed to the stronger effects that may be observed as the native state is approached.

Discussion

A weak native bias at the early stages of folding

The Levinthal paradox underscores how the folding of a protein into its native conformation

Table 1 Thermodynamic parameters and Φ-values for the early (β_T = 0.45), intermediate (β_T = 0.65), and late (β_T = 0.85) stages of folding

β_T = 0.45

Variant	$\Delta\Delta G_{\#}$ (kcal/mol)	$\Delta\Delta G_{D\text{-}N}$ (kcal/mol)	Φ
L4A	0.28 ± 0.03	1.9 ± 0.2	0.15 ± 0.01
T7S	0.05 ± 0.01	−0.08 ± 0.01	*
T8S	0.32 ± 0.03	0.77 ± 0.08	0.42 ± 0.05
L12S	0.21 ± 0.02	0.61 ± 0.06	0.34 ± 0.03
T16A	0.31 ± 0.03	2.6 ± 0.3	0.12 ± 0.01
L17A	0.55 ± 0.06	5.6 ± 0.6	0.10 ± 0.01
L20A	1.7 ± 0.2	6.5 ± 0.6	0.26 ± 0.03
L27A	0.72 ± 0.07	3.0 ± 0.3	0.24 ± 0.02
T33S	0.35 ± 0.04	1.2 ± 0.1	0.31 ± 0.03
V39A	0.51 ± 0.05	2.0 ± 0.2	0.26 ± 0.03
V45A	0.48 ± 0.05	1.6 ± 0.2	0.30 ± 0.03
L46A	0.84 ± 0.08	5.0 ± 0.5	0.17 ± 0.02
T47S	0.52 ± 0.05	1.1 ± 0.1	0.46 ± 0.05
V48A	1.1 ± 0.1	3.3 ± 0.3	0.34 ± 0.03
L50A	1.1 ± 0.1	4.7 ± 0.5	0.23 ± 0.02
T56S	0.43 ± 0.04	0.9 ± 0.1	0.48 ± 0.05
V58A	1.0 ± 0.1	2.0 ± 0.1	0.47 ± 0.05
T63S	0.59 ± 0.06	1.7 ± 0.2	0.35 ± 0.04
L70A	1.3 ± 0.3	1.6 ± 0.2	0.79 ± 0.06
T83S	0.48 ± 0.05	0.5 ± 0.1	0.92 ± 0.09
V88A	0.50 ± 0.05	1.4 ± 0.1	0.36 ± 0.03
V94A	0.48 ± 0.05	1.7 ± 0.2	0.28 ± 0.03
L96A	1.6 ± 0.2	5.6 ± 0.6	0.29 ± 0.03
L99A	−0.06 ± 0.01	1.8 ± 0.2	−0.03 ± 0.01
L100A	2.5 ± 0.3	7.6 ± 0.8	0.33 ± 0.03
L104A	1.0 ± 0.1	5.4 ± 0.5	0.19 ± 0.02
T105S	0.11 ± 0.01	−0.10 ± 0.01	*
L109A	0.14 ± 0.01	4.1 ± 0.4	0.03 ± 0.01
T110S	0.33 ± 0.03	1.0 ± 0.1	0.32 ± 0.03
L112A	0.12 ± 0.01	2.4 ± 0.2	0.05 ± 0.01
L114A	0.05 ± 0.01	3.4 ± 0.3	0.02 ± 0.01
L117A	0.15 ± 0.01	2.0 ± 0.2	0.08 ± 0.01

β_T = 0.65

Variant	$\Delta\Delta G_{\#}$ (kcal/mol)	$\Delta\Delta G_{D\text{-}N}$ (kcal/mol)	Φ
L4A	0.11 ± 0.01	1.9 ± 0.2	0.06 ± 0.01
T7S	−0.09 ± 0.01	−0.08 ± 0.01	*
T8S	0.12 ± 0.01	0.77 ± 0.08	0.15 ± 0.02
L12S	0.19 ± 0.02	0.61 ± 0.06	0.31 ± 0.03
T16A	−0.11 ± 0.01	2.6 ± 0.3	−0.04 ± 0.01
L17A	0.95 ± 0.05	5.6 ± 0.6	0.17 ± 0.01
L20A	3.2 ± 0.3	6.5 ± 0.7	0.50 ± 0.05
L27A	0.57 ± 0.06	3.0 ± 0.3	0.19 ± 0.01
T33S	−0.12 ± 0.01	1.1 ± 0.1	−0.11 ± 0.01
V39A	0.29 ± 0.04	2.0 ± 0.2	0.15 ± 0.01
V45A	0.66 ± 0.05	1.6 ± 0.2	0.41 ± 0.04
L46A	1.3 ± 0.1	5.0 ± 0.5	0.26 ± 0.03
T47S	0.70 ± 0.07	1.1 ± 0.1	0.62 ± 0.06
V48A	1.8 ± 0.2	3.3 ± 0.3	0.55 ± 0.05
L50A	1.7 ± 0.2	4.7 ± 0.5	0.35 ± 0.04
T56S	0.25 ± 0.02	0.89 ± 0.05	0.29 ± 0.03
V58A	0.91 ± 0.09	2.1 ± 0.2	0.44 ± 0.04
T63S	0.59 ± 0.06	1.7 ± 0.2	0.35 ± 0.04
L70A	1.1 ± 0.1	1.6 ± 0.2	0.67 ± 0.07
T83S	0.50 ± 0.05	0.52 ± 0.05	0.95 ± 0.10
V88A	0.63 ± 0.06	1.4 ± 0.1	0.46 ± 0.05
V94A	0.44 ± 0.04	1.7 ± 0.2	0.26 ± 0.03
L96A	2.3 ± 0.2	5.6 ± 0.6	0.41 ± 0.04
L99A	0.11 ± 0.01	1.8 ± 0.2	0.06 ± 0.01

(*continued on next page*)

Table 1 (*continued*)

β_T = 0.65			
Variant	**$\Delta\Delta G_{\#}$ (kcal/mol)**	**$\Delta\Delta G_{D\text{-}N}$ (kcal/mol)**	**Φ**
L100A	2.9 ± 0.3	7.6 ± 0.8	0.38 ± 0.03
L104A	1.9 ± 0.2	5.4 ± 0.5	0.35 ± 0.03
T105S	−0.07 ± 0.01	−0.12 ± 0.01	*
L109A	0.35 ± 0.04	4.1 ± 0.4	0.09 ± 0.01
T110S	−0.10 ± 0.01	1.0 ± 0.1	−0.10 ± 0.01
L112A	−0.18 ± 0.02	2.4 ± 0.2	−0.08 ± 0.01
L114A	−0.02 ± 0.01	3.4 ± 0.3	−0.01 ± 0.01
L117A	−0.16 ± 0.02	2.0 ± 0.2	−0.08 ± 0.01
β_T = 0.85			
Variant	**$\Delta\Delta G_{\#}$ (kcal/mol)**	**$\Delta\Delta G_{D\text{-}N}$ (kcal/mol)**	**Φ**
L4A	0.39 ± 0.04	1.9 ± 0.2	0.20 ± 0.02
T7S	−0.26 ± 0.03	−0.08 ± 0.01	*
T8S	0.01 ± 0.01	0.77 ± 0.08	0.01 ± 0.01
L12S	0.41 ± 0.04	0.61 ± 0.06	0.68 ± 0.07
T16A	0.12 ± 0.01	2.6 ± 0.3	0.05 ± 0.01
L17A	3.0 ± 0.3	5.6 ± 0.6	0.54 ± 0.05
L20A	3.7 ± 0.4	6.4 ± 0.6	0.56 ± 0.05
L27A	0.74 ± 0.07	3.0 ± 0.3	0.24 ± 0.02
T33S	−0.22 ± 0.02	1.1 ± 0.1	−0.19 ± 0.02
V39A	0.33 ± 0.03	2.0 ± 0.2	0.16 ± 0.02
V45A	0.71 ± 0.07	1.6 ± 0.2	0.43 ± 0.04
L46A	3.1 ± 0.3	5.0 ± 0.5	0.63 ± 0.06
T47S	0.63 ± 0.06	1.1 ± 0.1	0.55 ± 0.06
V48A	2.5 ± 0.3	3.3 ± 0.3	0.75 ± 0.08
L50A	1.8 ± 0.2	4.7 ± 0.5	0.39 ± 0.04
T56S	0.22 ± 0.02	0.89 ± 0.09	0.25 ± 0.03
V58A	1.0 ± 0.1	2.1 ± 0.2	0.50 ± 0.05
T63S	0.72 ± 0.07	1.7 ± 0.2	0.43 ± 0.04
L70A	0.86 ± 0.09	1.6 ± 0.2	0.54 ± 0.05
T83S	0.19 ± 0.02	0.52 ± 0.05	0.36 ± 0.04
V88A	0.65 ± 0.07	1.4 ± 0.1	0.47 ± 0.05
V94A	0.37 ± 0.04	1.7 ± 0.2	0.22 ± 0.02
L96A	1.7 ± 0.2	5.6 ± 0.6	0.31 ± 0.03
L99A	0.10 ± 0.01	1.8 ± 0.2	0.06 ± 0.01
L100A	2.9 ± 0.3	7.6 ± 0.8	0.39 ± 0.04
L104A	2.3 ± 0.2	5.4 ± 0.5	0.42 ± 0.04
T105S	0.20 ± 0.02	−0.12 ± 0.01	*
L109A	0.73 ± 0.07	4.1 ± 0.4	0.18 ± 0.02
T110S	−0.12 ± 0.01	1.0 ± 0.1	−0.12 ± 0.01
L112A	−0.09 ± 0.01	2.4 ± 0.2	−0.04 ± 0.01
L114A	0.57 ± 0.06	3.4 ± 0.3	0.17 ± 0.02
L117A	−0.04 ± 0.01	2.0 ± 0.2	−0.02 ± 0.01

*The variant produced a change in stability that was too low to determine a reliable value of Φ.

does not take place through a random exploration of all possible states.[31] One of the most elegant approaches to solve this paradox involves postulating that proteins are characterized by a funnel-shaped energy landscape, implying the existence of minimal frustration.[18,19,32]

In the case of a minimally frustrated sequences, a small set of collective parameters that gauge proximity to the native arrangement can already serve as potent indicators of a given configuration energy.[33] Consequently, the ultimate conformation of the protein exerts a profound influence on the topology of the folding landscape. This results in entire family of proteins, which ultimately adopt the same structure, sharing analogous landscapes.[2,4,5,34–36] Because chain topology has a dominant effect on the energy landscape of folding, it follows that sequence variations between homologous proteins are likely to affect local frustration patterns on the most heterogeneous states along the pathway. Hence, the manifestation of such hypothesis implies that sequence variations between homologues may produce little changes which are confined in the early stages of folding,[2,37] where more dynamic conformations are likely to be explored.

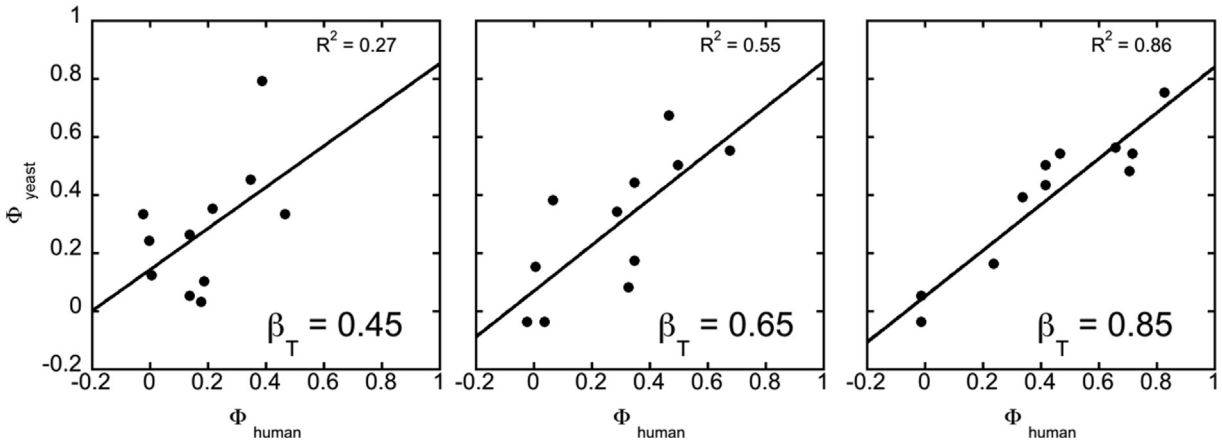

Figure 4. Φ-Φ comparison for yeast and human frataxin. At higher β_T the correlation of Φ-values conforms to a straight line with a slope of 1 highlighting a high robustness of the folding reaction, while a weaker native bias is reported in the earlier events of folding.

Table 2 Comparison of the Φ values obtained at homologous positions for yeast and human frataxin

β_T = 0.45

Yeast variant	Δ Δ G# (kcal/mol)	Human variant	Φ
Y29A	0.01 ± 0.04	T16A	0.12 ± 0.01
L30A	0.19 ± 0.03	L17A	0.10 ± 0.01
L33A	0.14 ± 0.01	L20A	0.26 ± 0.03
V51A	0 ± 0.01	V39A	0.26 ± 0.03
L60A	0.47 ± 0.02	V48A	0.34 ± 0.03
I70V	0.35 ± 0.01	V58A	0.47 ± 0.05
I79V	0.22 ± 0.02	T63S	0.35 ± 0.04
L81A	0.39 ± 0.03	L70A	0.79 ± 0.06
L107A	−0.02 ± 0.01	L100A	0.33 ± 0.03
V115A	0.07 ± 0.01	L109A	0.03 ± 0.01
I119V	0.14 ± 0.01	L112A	0.05 ± 0.01

β_T = 0.65

Yeast variant	Δ Δ G# (kcal/mol)	Human variant	Φ
Y29A	−0.02 ± 0.05	T16A	−0.04 ± 0.01
L30A	0.35 ± 0.09	L17A	0.17 ± 0.01
L33A	0.50 ± 0.02	L20A	0.50 ± 0.05
V51A	0.01 ± 0.01	V39A	0.15 ± 0.01
L60A	0.68 ± 0.03	V48A	0.55 ± 0.05
I70V	0.35 ± 0.02	V58A	0.44 ± 0.04
I79V	0.29 ± 0.02	T63S	0.35 ± 0.04
L81A	0.47 ± 0.09	L70A	0.67 ± 0.07
L107A	0.07 ± 0.03	L100A	0.38 ± 0.03
V115A	0.33 ± 0.04	L109A	0.09 ± 0.01
I119V	0.04 ± 0.3	L112A	−0.08 ± 0.01

β_T = 0.85

Yeast variant	Δ Δ G# (kcal/mol)	Δ Δ G_{D-N} (kcal/mol)	Φ
Y29A	−0.01 ± 0.05	T16A	0.05 ± 0.01
L30A	0.72 ± 0.24	L17A	0.54 ± 0.05
L33A	0.66 ± 0.05	L20A	0.56 ± 0.05
V51A	0.24 ± 0.10	V39A	0.16 ± 0.02
L60A	0.83 ± 0.06	V48A	0.75 ± 0.08
I70V	0.42 ± 0.18	V58A	0.50 ± 0.05
I79V	0.42 ± 0.06	T63S	0.43 ± 0.04
L81A	0.47 ± 0.18	L70A	0.54 ± 0.05
L107A	0.34 ± 0.1	L100A	0.39 ± 0.04
V115A	0.71 ± 0.05	L109A	0.18 ± 0.02
I119V	−0.1 ± 0.04	L112A	−0.04 ± 0.01

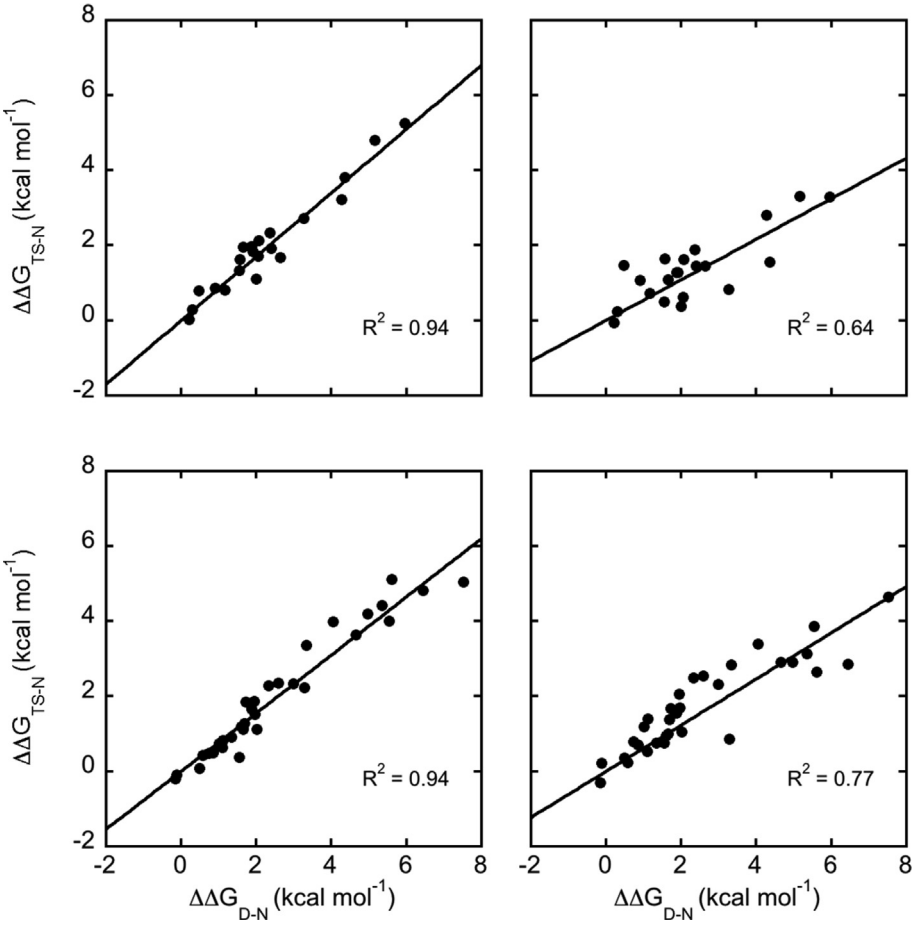

Figure 5. Brønsted plots of the early (left panels) and late (right panels) events of folding of both yeast (top) and human (bottom) frataxin. Lines represent the best fit to a straight line.

It is of particular interest to discuss the comparative folding analysis reported in this work on the light of the predictions ensuing from the minimal frustration theory outlined above. In fact, in line with expectations, the two homologous frataxins fold via a similar mechanism. However, a closer look based on the mutational analysis of the early, intermediate and late events of folding clearly shows a more complex picture, with a weaker native bias at the early stages of folding. This finding thus represents an indirect experimental validation of the energy landscape theory of protein folding and, at the same time, offers a glimpse of the width of such landscape, exemplifying the effects of sequence variations on the early stages of the reaction.

'**Nucleation**' **or** '**condensation**' **– what comes first?**

The possibility to infer both early and late events of folding of frataxins makes the comparative analysis particularly instructive also when describing them on the light of previously identified mechanisms of folding. Several different protein systems have been described to fold via a general mechanism, classically denoted as nucleation-condensation.[38–41] The model postulates the formation of an extended nucleus stabilized by a general compaction of the whole polypeptide chain. An important implication of nucleation-condensation is that formation of the nucleus (nucleation) is coupled with a more general organization of structure (condensation), the transition state resembling a distorted version of the native state. A characteristic signature of the nucleation-condensation scenario may be found when studying the effect of structural perturbation on the activation and ground state free energies (Brønsted plot analysis).[40,41] In fact, linearity in a Brønsted profile is indicative of a general compaction of the polypeptide chain with native-like structure homogeneously distributed in the whole globule. Conversely, the presence of scatter in the Brønsted may be associated with discrete elements of structure that may form independently from the reminder of the macromolecule.[4]

But is it possible to detect a decoupling between the nucleation and condensation steps? Figure 5 reports the Brønsted profile of the early and late events of folding of both human and yeast

frataxin. It is evident that in all cases there is a linear correlation between the changes in activation energies with those of the ground states characterized by different thermodynamic stabilities. Remarkably, however, in both proteins scatter increases when approaching the native state, with an R^2 going from 0.94 at β_T = 0.45, for both proteins, to R^2 = 0.64 and R^2 = 0.77 at β_T = 0.85 for yeast and human frataxin respectively. This finding suggests that, whilst the transition state displays native-like properties both in the early and late events, in the case of both proteins the global compaction of the polypeptide chain appears to precede the formation of discrete and structurally localized clusters of specific interactions.

Conclusions

The study of protein systems displaying a broad energy barrier for folding offers the appealing opportunity to address the properties of early, intermediate and late events. Furthermore, the comparative analysis of homologous proteins as exemplified in this work, represents a valuable tool to extract general rules of such a complex reaction. This opportunity complements the study of folding transition paths, which can be obtained from single molecule approaches.[42,43] In the case of frataxin, the comparison between the human and yeast homologues allows discovering a weak native bias at the early stages of folding, with observed pathways converging only when the native state is approached. Additionally, it appears that consolidation of the folding nucleus is slightly decoupled from a general collapse of the polypeptide chain, which appears to precede nucleation. Future comparative analysis on other systems characterized by a broad energy barrier will be critical to further reinforce to general nature of such observations.

Materials and Methods

Site-directed mutagenesis

Human frataxin gene was subcloned in a pET28b + plasmid vector. The constructs encoding the site-directed variants of human frataxin were obtained using the gene encoding for human frataxin wt as a template to perform site-directed mutagenesis using QuikChange® Lightning kit from Agilent Technologies, according to the instructions of the manufacturer. The presence of the desired mutations and the absence of unwanted ones were confirmed by sequence analysis.

Protein expression and purification

Human frataxin wild type and all the site-directed variants were expressed in E. coli cells BL21 (DE3). Bacterial cells were grown at 37 °C in LB medium containing kanamycin as an antibiotic at a final concentration of 30 μg/ml until optical density OD600 reached 0.6. The protein expression was induced overnight by adding 0.5 mM IPTG (isopropyl-β-D-thiogalactoside). After induction, cells were grown overnight at 18 °C and then collected by centrifugation. To purify the His-tagged protein, the pellet was resuspended in 40 ml of binding buffer (20 mM Tris- HCl, 500 mM NaCl, 5 mM Imidazole, pH 8.0) containing 1 mM TCEP (Tris(2-carboxyethyl) phosphine) in the presence of antiprotease tablet (Complete EDTA-free, Roche), then sonicated and centrifuged. The soluble fraction from bacterial lysate was loaded onto a nickel-charged HisTrap Chelating HP (GE Healtcare) column equilibrated with the resuspending buffer. Protein was then eluted with a gradient from 0 to 250 mM Imidazole using an ÄKTA prime system. Fractions containing the protein were collected, and the buffer was exchanged to 20 mM Tris-HCl, 500 mM NaCl, 1 mM TCEP, pH 8.0 using a HiTrap Desalting column (GE Healthcare). The purity of the protein was analyzed through on a pre-casted NuPage 4–12 % Bis-Tris polyacrylamide gel (Invitrogen, Carlsbad, CA). Protein concentration was estimated by measuring the absorbance of tryptophan residue at 280 nm and calculated through the lambert-Beer equation.

Stopped – Flow (un)Folding kinetics experiments

Kinetic (un)folding experiments were carried out on an Applied Photophysics Pi-star 180 stopped-flow instrument, monitoring the change of fluorescence emission; the excitation wavelength was 280 nm and the fluorescence emission was collected using a 320 nm cut-off glass filter for unfolding experiments and a 360-nm cut-off glass filter for refolding experiments. The experiments were performed at 37 °C in 50 mM Tris-HCl, 100 mM NaCl, 1 mM DTT, pH 8.0, by using urea as denaturant. At least five individual traces were acquired and then averaged for each denaturant concentration. Protein concentration was typically 2 μM. In all cases the fluorescence time courses obtained was satisfactorily fitted y using a single exponential equation.

Funding

This work was partly supported by grants from Sapienza University of Rome (RP11715C34AEAC9B, RM1181641C2C24B9, RM11916B414C897E, RG12017297FA7223 to S. G., RM12218148DA1933 to A.T.), by an ACIP grant (ACIP 485–21) from Institut Pasteur Paris to S.G., the Associazione Italiana per la Ricerca sul Cancro (Individual Grant–IG 24551 to S.G.), the Istituto Pasteur Italia (“Teresa Ariaudo Research

Project" 2018, and "Research Program 2022 to 2023 Under 45 Call 2020" to A.T.), the Italian MUR-PRIN 2022 grant N. 2022JY3PMB to A.T., the European Union's Horizon 2020 Research and Innovation program under the Marie Skłodowska–Curie Grant Agreement UBIMOTIF No 860517 (to S.G.). 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.

CRediT authorship contribution statement

Paola Pietrangeli: Writing – review & editing, Investigation, Formal analysis. **Lucia Marcocci:** Writing – review & editing, Investigation, Formal analysis. **Valeria Pennacchietti:** Investigation, Formal analysis. **Awa Diop:** Investigation, Formal analysis. **Mariana Di Felice:** Investigation, Formal analysis. **Livia Pagano:** Investigation, Formal analysis. **Francesca Malagrinò:** Writing – review & editing, Investigation, Formal analysis. **Angelo Toto:** Writing – review & editing, Investigation, Formal analysis, Conceptualization. **Maurizio Brunori:** Writing – review & editing, Conceptualization. **Stefano Gianni:** Writing – review & editing, Writing – original draft, Supervision, Conceptualization.

DECLARATION OF COMPETING INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Received 1 February 2024;
Accepted 25 March 2024;
Available online 27 March 2024

Keywords:
protein folding;
kinetics;
mutagenesis;
energy landscape;
linear-free energy relationships

† These authors contributed equally to this work.

References

1. Fersht, A.R., (1999). Structure and Mechanism in Protein Science. Freeman, New York.
2. Calosci, N., Chi, C.N., Richter, B., Camilloni, C., Engstrom, Å., Eklund, L., Travaglini-Allocatelli, C., Gianni, S., Vendruscolo, M., Jemth, P., (2008). Comparison of successive transition states for folding reveals alternative early folding pathways of two homologous proteins. *Proc. Natl. Acad. Sci. U S A* **105**, 19241–19246.
3. Clarke, J., Cota, E., Fowler, S.B., Hamill, S.J., (1999). Folding studies of Ig-like beta-sandwich proteins suggest they share a common folding pathway. *Structure* **7**, 1145–1153.
4. Gianni, S., Guydosh, N.R., Khan, F., Caldas, T.D., Mayor, U., White, G.W., DeMarco, M.L., Daggett, V., Fersht, A.R., (2003). Unifying features in protein-folding mechanisms. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 13286–13291.
5. Scott, K.A., Randles, L.G., Clarke, J., (2004). The folding of spectrin domains II: phi-value analysis of R16. *J. Mol. Biol.* **344**, 207–211.
6. Travaglini-Allocatelli, C., Gianni, S., Morea, V., Tramontano, A., Soulimane, T., Brunori, M., (2003). Exploring the cytochrome c folding mechanism: cytochrome c552 from thermus thermophilus folds through an on-pathway intermediate. *J. Biol. Chem.* **278**, 41136–41140.
7. Zarrine-Afsar, A., Larson, S.M., Davidson, A.R., (2005). The family feud: do proteins with similar structures fold via the same pathway?. *Curr. Opin. Struct. Biol.* **15**, 42–49.
8. Adinolfi, S., Iannuzzi, C., Prischi, F., Pastore, C., Iametti, S., Martin, S.R., Bonomi, F., Pastore, A., (2009). Bacterial frataxin CyaY is the gatekeeper of iron-sulfur cluster formation catalyzed by IscS. *Nature Struct. Mol. Biol.* **16**, 390–396.
9. Adinolfi, S., Trifuoggi, M., Politou, A.S., Martin, S., Pastore, A., (2002). A structural approach to understanding the iron-binding properties of phylogenetically different frataxins. *Hum. Mol. Genet.* **11**, 1865–1877.
10. Pandolfo, M., Pastore, A., (2009). The pathogenesis of Friedreich ataxia and the structure and function of frataxin. *J. Neurol.* **256**, 9–17.
11. Pastore, A., Puccio, H., (2013). Frataxin: a protein in search for a function. *J. Neurochem.* **126**, 43–52.
12. Pastore, C., Franzese, M., Sica, F., Temussi, P., Pastore, A., (2007). Understanding the binding properties of an unusual metal-binding protein–a study of bacterial frataxin. *FEBS J.* **274**, 4199–4210.
13. Bonetti, D., Toto, A., Giri, R., Morrone, A., Sanfelice, D., Pastore, A., Temussi, P.A., Gianni, S., Brunori, M., (2014). The kinetics of folding of frataxin. *Phys. Chem. Chem Phys.*
14. Gianni, S., Camilloni, C., Giri, R., Toto, A., Bonetti, D., Morrone, A., Sormanni, P., Brunori, M., Vendruscolo, M., (2014). Understanding the frustration arising from the competition between function, misfolding, and aggregation in a globular protein. *Proc. Natl. Acad. Sci. USA* **111**, 14141–14146.
15. Otzen, D.E., Kristensen, O., Proctor, M., Oliveberg, M., (1999). Structural changes in the transition state of protein folding: alternative interpretations of curved chevron plots. *Biochemistry* **38**, 6499–6511.

16. Ternstrom, T., Mayor, U., Akke, M., Oliveberg, M., (1999). From snapshot to movie: phi analysis of protein folding transition states taken one step further. *Proc. Natl. Acad. Sci. U S A* **96**, 14854–14859.
17. Sanfelice, D., Morandi, E., Pastore, A., Niccolai, N., Temussi, P.A., (2015). Cold denaturation unveiled: molecular mechanism of the asymmetric unfolding of yeast frataxin. *Chemphyschem* **16**, 3599–3602.
18. Bryngelson, J.D., Onuchic, J.N., Socci, N.D., Wolynes, P. G., (1995). Funnels, pathways, and the energy landscape of protein folding: a synthesis. *Proteins* **21**, 167–195.
19. Onuchic, J.N., Socci, N.D., Luthey-Schulten, Z., Wolynes, P.G., (1996). Protein folding funnels: the nature of the transition state ensemble. *Fold. Des.* **1**, 441–450.
20. Fersht, A.R., 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.
21. Fersht, A.R., Sato, S., (2004). Phi-value analysis and the nature of protein-folding transition states. *Proc. Natl. Acad. Sci. U S A* **101**, 7976–7981.
22. Khan, F., Chuang, J.I., Gianni, S., Fersht, A.R., (2003). The kinetic pathway of folding of barnase. *J. Mol. Biol.* **333**, 169–186.
23. Khorasanizadeh, S., Peters, I.D., Roder, H., (1996). Evidence for a three-state model of protein folding from kinetic analysis of ubiquitin variants with altered core residues. *Nature Struct. Biol.* **3**, 193–205.
24. Parker, M.J., Spencer, J., Clarke, A.R., (1995). An integrated kinetic analysis of intermediates and transition states in protein folding reactions. *J. Mol. Biol.* **253**, 771–786.
25. Sanchez, I.E., Kiefhaber, T., (2003). Evidence for sequential barriers and obligatory intermediates in apparent two-state protein folding. *J. Mol. Biol.* **325**, 367–376.
26. Oliveberg, M., (1998). Alternative explanations for multi-state kinetics in protein folding: transient aggregation and changing transition-state ensambles. *Acc. Chem. Res.* **31**, 765–772.
27. Oliveberg, M., (2001). Characterisation of the transition states for protein folding: towards a new level of mechanistic detail in protein engineering analysis. *Curr. Opin. Struct. Biol.* **11**, 94–100.
28. Gianni, S., Brunori, M., Jemth, P., Oliveberg, M., Zhang, M., (2009). Distinguishing between smooth and rough free energy barriers in protein folding. *Biochemistry* **48**, 1825–1830.
29. Hammond, G.S., (1955). A correlation of reaction rates. *J. Am. Chem. Soc.* **77**, 334–339.
30. Gianni, S., Jemth, P., (2014). Conserved nucleation sites reinforce the significance of Phi value analysis in protein-folding studies. *IUBMB Life* **66**, 449–452.
31. Levinthal, C., (1968). Are there pathways for protein folding?. *J. Chem. Phys.* **65**, 44–45.
32. Wolynes, P.G., (2005). Energy landscapes and solved protein-folding problems. *Philos. Transact. R. Soc. A Math. Phys. Eng. Sci.* **363**, 453–464.
33. Oliveberg, M., Wolynes, P.G., (2005). The experimental survey of protein-folding energy landscapes. *Q. Rev. Biophys.* **38**, 245–288.
34. Hultqvist, G., Pedersen, S.W., Chi, C.N., Strømgaard, K., Gianni, S., Jemth, P., (2012). An expanded view of the protein folding landscape of PDZ domains. *Biochem. Biophys. Res. Commum.* **421**, 550–553.
35. Ivarsson, Y., Travaglini-Allocatelli, C., Brunori, M., Gianni, S., (2008). Folding and misfolding in a naturally occurring circularly permuted PDZ domain. *J. Biol. Chem.* **283**, 8954–8960.
36. Steward, A., Chen, Q., Chapman, R.I., Borgia, M.B., Rogers, J.M., Wojtala, A., Wilmanns, M., Clarke, J., (2012). Two immunoglobulin tandem proteins with a linking β-strand reveal unexpected differences in cooperativity and folding pathways. *J. Mol. Biol.* **416**, 137–147.
37. Di Silvio, E., Brunori, M., Gianni, S., (2015). Frustration sculpts the early stages of protein folding. *Angew. Che. Int. Ed. Engl.* **54**
38. Abkevich, V.I., Gutin, A.M., Shakhnovich, E.I., (1994). Specific nucleus as the transition state for protein folding: evidence from the lattice model. *Biochemistry* **33**, 10026–10036.
39. Fersht, A.R., (1995). Optimization of rates of protein folding: the nucleation-condensation mechanism and its implications. *Proc. Natl. Acad. Sci. U. S. A.* **21**, 10869–10873.
40. Fersht, A.R., (2004). Relationship of Leffler (Bronsted) alpha values and protein folding Phi values to position of transition-state structures on reaction coordinates. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 14338–14342.
41. Itzhaki, L.S., Otzen, D.E., Fersht, A.R., (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.
42. Chung, H.S., Eaton, W.A., (2013). Single-molecule fluorescence probes dynamics of barrier crossing. *Nature* **502**, 685–688.
43. Chung, H.S., Eaton, W.A., (2018). Protein folding transition path times from single molecule FRET. *Curr. Opin. Struct. Biol.* **48**, 30–39.