

Published in final edited form as:

Nature. 2010 February 4; 463(7281): 685–688. doi:10.1038/nature08743.

Experimental evidence for a frustrated energy landscape in a 3-helix bundle protein family

Beth G. Wensley, Sarah Batey[†], Fleur A. C. Bone, Zheng Ming Chan, Nuala R. Tumelty, Annette Steward, Lee Gyan Kwa, Alessandro Borgia[†], and Jane Clarke

Department of Chemistry, University of Cambridge, MRC Centre for Protein Engineering, Lensfield Rd, Cambridge CB2 1EW UK.

Abstract

Energy landscape theory is a powerful tool for understanding the structure and dynamics of complex molecular systems, in particular biological macromolecules¹. The primary sequence of a protein defines its free energy landscape, and thus determines the folding pathway and the rate constants of folding and unfolding, as well as its native structure. Theory has shown that roughness in the energy landscape will lead to slower folding¹, but derivation of detailed experimental descriptions of this landscape is challenging. Simple folding models^{2,3} show that folding is significantly influenced by chain entropy; proteins where the contacts are local fold fast, due to the low entropy cost of forming stabilising, native contacts during folding^{4,5}. For some protein families, stability is also a determinant of folding rate constants⁶. Where these simple metrics fail to predict folding behaviour it is probable that there are features in the energy landscape that are unusual. Such general observations cannot explain the folding behaviour of the R15, R16 and R17 domains of α -spectrin. R15 folds ~3000 times faster than its homologues, although they have similar structures, stabilities and, as far as can be determined, transition state stabilities⁷⁻¹⁰. Here we show that landscape roughness (internal friction) is responsible for the slower folding and unfolding of R16 and R17. We use chimeric domains to demonstrate that this internal friction is a property of the core, and suggest that frustration in the landscape of the slow folding spectrin domains may be due to mis-docking of the long helices during folding. Although theoretical studies have suggested that rugged landscapes will result in slower folding, this is the first time that such a phenomenon has been shown experimentally to directly influence the folding kinetics of a “normal” protein with a significant energy barrier – one which folds on a relatively slow ms-s timescale.

The folding rate constants of all but the fastest folding proteins are assumed to be determined by the free energy barrier between the unfolded and transition states and a kinetic pre-factor. This is reflected in the ability of simple structural parameters such as contact or long-range order^{2,3}, to predict folding rate constants. R15, R16 and R17 (~30% pairwise sequence identity) have the same three-helix bundle structure (Fig. 1a), and similar thermodynamic stabilities, but R15 folds and unfolds ~3 orders-of-magnitude faster than R16 and R17 (Fig. 2a)⁷. The folding rate constant for R15 is well-predicted by contact and long range order plots but R16 and R17 are outliers (Supplementary Fig. 1). Since the

Correspondence and requests for materials should be addressed to J.C. (jc162@cam.ac.uk).

[†]Current addresses: SB: Covagen AG, c/o ETH Zürich, Y17 M22, Life Science ETH/Uni, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland; AB: Biochemisches Institut, Universität Zürich, Winterthurerstrasse 190 CH-8057, Zürich, Switzerland.

Author Contributions: B.G.W., S.B. and J.C. designed the investigation. B.G.W., S.B., F.A.C.B., Z.M.C., N.R.T., A.S. and L.G.K. performed the experiments and B.G.W. and S.B. did most of the analysis. A.B. contributed to discussions. B.G.W. and J.C. wrote the paper. The authors declare no competing financial interests.

Supplementary Information accompanies this paper

difference is evident in both folding and unfolding, the simplest explanation is that the transition state (TS) of R15 is more structured (and thus more stable) than that of its homologues. It is not possible to measure the free energy of the TS directly, however experimental data show that the three TSs are similar in terms of compactness (β_T), and protein engineering Φ -value analysis of all three suggests that they are generally similar in overall structural and energetic terms (same average Φ -values, and same regions of the structure folded)⁸⁻¹⁰. If such significant differences in folding kinetics cannot easily be ascribed to differences in TS structure it is necessary to look for an alternative explanation.

Energy landscape theory has shown that protein folding is best described as a Kramers'-like¹¹ diffusive process across the energy landscape and the folding rate constant is dependant on a number of aspects of the landscape¹. Explicitly, for a one-dimensional free-energy surface with harmonic wells, the folding time (τ_f) is related to the shape of the

energy landscape ($\omega_p, \omega_{\ddagger}$), the height of the energy barrier (ΔG^\ddagger) and the diffusion coefficient (D_0);

$$\tau_f = \frac{1}{k_f} = \frac{2\pi RT}{\omega_p \omega_{\ddagger} D_0} \exp\left(\frac{\Delta G^\ddagger}{RT}\right) \quad (1)$$

One principle component of the diffusion coefficient is solvent friction, which slows folding. Diffusion over smooth landscapes is relatively fast, so that folding time of a given protein generally depends only on the height of the free energy barrier and solvent viscosity. However, theory suggests that when the landscape is rough, diffusion will be slowed by the time taken for the chain to escape from the local minima which constitute the rougher (more frustrated) landscape¹. One possible explanation for the slow folding of R16 and R17 is that diffusion is impeded by kinetic traps in the energy landscape. This would be remarkable because landscape roughness has never been observed explicitly for any two-state protein with a significant energy barrier (i.e. a relatively slow folding protein with folding times in the order of ms – s), although frustration resulting from formation of stable misfolded intermediates has been described¹².

A simple way to test this hypothesis is to determine the dependence of folding and unfolding rate constants on solvent viscosity^{13,14}. Equation 1 has been recast, empirically¹⁵, as:

$$k = \frac{C}{\eta + \sigma} \cdot \exp\left(\frac{-\Delta G^\ddagger}{RT}\right) \quad (2)$$

to relate the rate constant for folding (or unfolding) (k_f, k_u) to solvent viscosity, η , and the so-termed internal friction of the protein, σ . Thus for a system with a smooth energy landscape (where internal friction is negligible i.e. $\sigma \ll \eta$), k_f will be inversely proportional to solvent viscosity. Such a relationship has been observed for a number of small proteins¹⁶⁻¹⁸¹.

Solvent viscosity is easily controlled through addition of small molecule viscosogens, however these tend to increase the stability of proteins. The isostability approach, whereby the stabilisation caused by the viscogen is counteracted by the addition of a chemical denaturant, has been widely used^{16-18,20,21}. We applied this approach to investigate the

¹Solvent slaving has been proposed as an alternative mechanism to explain the effects of solvent viscosity on folding kinetics¹⁹. This is discussed and compared to the model used in this work in the Supplementary Results and Discussion.

hypothesis that internal friction slows the folding of R16 and R17, but is insignificant in the fast folding R15.

The equilibrium stability and folding kinetics of all three proteins were investigated over a range of solvent viscosities using glucose as the viscogen, and guanidinium chloride (GdmCl) as denaturant (Supplementary Figs 2 and 3). Glucose increases the stability of all the proteins but the m -value decreases. This decrease in m -value is associated with a decrease in the refolding m -value, but the unfolding m -value is constant, suggesting that glucose causes a collapse of the denatured states (Supplementary Fig. 4). This effect is seen in all three proteins, and more importantly, since the unfolding m -value is unaffected by the viscogen this suggests that the TS position (and structure) is unaffected by glucose in all cases. Folding and unfolding rate constants were determined at isostability in two ways. (i) The equilibrium data were used to calculate the concentration of GdmCl at which $\Delta G_{D-N} = 1.5 \text{ kcal mol}^{-1}$ for each glucose concentration. k_f and k_u values at this GdmCl concentration were determined from the fits of the chevron plots. (ii) The chevron plots alone were used to determine rate constants at $\Delta G_{D-N} = 0$ (i.e. where $k_f = k_u$).

Relative rate constants (k_f/k_u) were plotted vs. relative viscosity (η/η_0) (Fig. 2b). Gradients for folding and unfolding rate constants were consistent. From Equation 2, at isostability, if σ is small relative to η , the gradient of this plot will be approximately 1. If, however, σ is large relative to η , the gradient will be significantly < 1 . The data for R15 show that both the folding and unfolding rate constants are strongly dependant on solvent viscosity, with gradients close to 1 (Supplementary Table 1), suggesting that for R15 internal friction is negligible. However, the rate constants of R16 and R17 show only a very weak dependence on solvent viscosity (mean slopes ~ 0.2). This provides strong evidence to suggest that the reason R16 and R17 fold so much more slowly than R15 is due, at least in part, to internal friction i.e. to roughness in the energy landscape. Similar investigations have been undertaken for four other proteins which fold on timescales comparable to those of the spectrin domains 16-18,20. As seen here for R15, there is no evidence to suggest that internal friction plays any role in determining the rate constant for folding of these proteins. R16 and R17 are unexpectedly atypical.

Using equation (2) it is possible to estimate the internal friction, σ , in the transition states of the spectrin domains. (See Supplementary Results/Discussion). R15 has a value of σ of $0.25 \pm 0.16 \text{ cP}$, significantly lower than water ($\eta \sim 1 \text{ cP}$). By contrast, the σ values for R16 and R17 are significantly higher than water ($4.4 \pm 1.6 \text{ cP}$, and $12.0 \pm 6.6 \text{ cP}$, respectively). These high values of internal friction are similar to those found in studies of dynamics of essentially fully folded proteins^{15,22}.

These values of σ can be used to evaluate the relative magnitude of the ruggedness of the energy landscapes of our spectrin domains. Assuming randomly (Gaussian) distributed roughness, Zwanzig²³ obtained an expression relating the amplitude of the roughness (characterised by variance ϵ^2) on a one-dimensional energy landscape to the effective diffusion coefficient, D^* :

$$D^* = D \exp \left[-(\epsilon/RT)^2 \right] \quad (3)$$

where D is the diffusion coefficient across a given smooth landscape and ϵ is the characteristic magnitude of the roughness (in energy units). This expression can be used to estimate the relative landscape roughness of the slow folding proteins R16 and R17, compared to R15, $\Delta\epsilon$ (see Supplementary Results/Discussion). $\Delta\epsilon$ is $\sim 1.7 RT$ for R16 and $\sim 2.0 RT$ for R17, similar to the value observed in experimental studies on peptides, denatured proteins and small, fast folding proteins and from theory²⁴⁻²⁸. As discussed by

Zwanzig, Gaussian noise gives one limit for the microscopic barrier heights; the alternative extreme of periodic barriers of equal height results in a slightly larger Δe . It is remarkable that such a small increase in landscape roughness can result in such a significant change in the viscosity dependence of the folding kinetics. In fact, these microscopic barriers must be relatively small; kinetic traps involving barriers much larger than about 3-4 RT would result in accumulation of intermediates, not observed in spectrin domains^{1,12,27}.

What then, is the source of the frustration in the energy landscape observed in R16 and R17? More than 200 variants of the spectrin domains have been investigated⁸⁻¹⁰. None significantly speeds the folding of R16 / R17 nor significantly slows the folding of R15. Four core-swap proteins, were designed (Fig. 1b,c and Supplementary Fig. 5) containing the core residues from one parent, the “minor parent”, grafted into the “major parent” which contributes the “outside” of the core-swapped protein. R15o16c (the outside of R15 and the core of R16) and R15o17c could not be expressed solubly. However, both proteins containing the core residues from R15, R16o15c and R17o15c, fold and unfold significantly faster than their major parent (R16 and R17), suggesting that the origin of fast/slow folding lies within the core (Figs. 3a & 3b).

Furthermore, for both R16o15c and, more notably, R17o15c the increase in the rate constants is accompanied by an increase in the dependence of the rate constants on solvent viscosity, compared to their major parent (Fig. 3c, Supplementary Fig. 6a,b, Supplementary Table 1). The slopes of the relative rate constant vs. relative viscosity plots for R17o15c are comparable to those of R15, and significantly different to its major parent R17. For R16o15c, the slopes are significantly higher than those of R16, but lower than R15 or R17o15c. Thus faster folding is associated with a decrease in internal friction.

Mechanistic differences may offer insight into the source of the roughness in the energy landscapes of the slow-folding spectrin domains. R16 and R17 fold via a framework-like mechanism, in which formation of the helices precedes helix packing^{8,9}. However, R15 folds by nucleation-condensation, where secondary and tertiary structure form concomitantly¹⁰. Evidence for different folding mechanisms is seen most clearly in the C helix (Fig. 4a). Φ -value analysis of the C helix of R16o15c clearly indicates that the pattern of Φ -values resembles R15, not R16 (Fig. 4, Supplementary Table 2, Supplementary Fig. 7, Supplementary Fig. 8). Thus slow folding and increased internal friction may be related to a framework-like folding mechanism. (Note that R16o15c has significantly reduced helical propensity in the C-helix, compared to R16 (Supplementary Fig. 9)).

In R15 nucleation, involving the central regions of the A and C helices, establishes the correct register for the docking of these long helices. In R16 and R17 however, the C helix (and to lesser extent A helix) is apparently pre-formed, and must find the correct register to dock. A potential source of conformational frustration is the occurrence of a number of non-native docking events as the polypeptide chain crosses the TS barrier. Indeed, early, out-of-register mis-docking events are seen in MD simulations of R17 folding⁹. We propose that this misfolding is a likely source of the frustration in the folding landscapes of R16 and R17, and note that transient mis-docking would be likely to result in roughness of the magnitude observed for transient contact formation in unfolded peptides (0.5-2 RT) as we find here. Paradoxically, studies of the 3-helix bundle homeodomain family of proteins suggest that folding via a framework mechanism results in faster folding than nucleation condensation²⁹. The difference between the spectrin and homeodomain proteins is the size of the helices. There is, perhaps, little scope for mis-docking in the small, ~12 residue homeodomain helices whereas for the long (~30 residue) spectrin helices it is more difficult to establish the correct alignment.

All our results are consistent with the hypothesis that the slow folding and unfolding of R16 and R17 are due to roughness in the energy landscapes. We suggest that this friction results from residue-specific phenomena, such as frustration caused by mis-docking of helices. Although theoretical studies have long suggested that frustrated landscapes will result in slower folding, this is the first time that such a phenomenon has been shown experimentally to directly influence the folding kinetics of a two-state protein which folds on a relatively slow ms - s timescale. It is possible that slow folding / unfolding kinetics might be advantageous to proteins, such as spectrins, which have very long half-lives *in vivo*. Spectrin is a protein of the intracellular matrix of red-blood cells where it is important for membrane elasticity. Red blood cells live for ~120 days. Slow unfolding kinetics will result in far fewer domain unfolding events during this lifetime, perhaps decreasing the likelihood of degradation, or other detrimental effects. We note that mutations in spectrin domains which reduce inter-domain cooperativity, and thus also increase the likelihood of domain unfolding, result in disease³⁰.

Methods Summary

Protein expression and purification was carried out as described elsewhere^{7,10}. Design of the core-swapped proteins is described in the **Methods**. Equilibrium stability was determined monitoring the CD at 222 nm, and kinetics followed changes in fluorescence as described in 7. Methods of fitting the kinetic data are described in detail in the **Methods**. Note that our previous work on these domains has been carried out in urea⁷⁻¹⁰, however due to the stabilising effect of the glucose, the stronger denaturant GdmCl was used. The exception was R17o15c: due to a combination of the destabilisation and the effect of ionic strength on its stability, all analysis of this domain was carried out in urea.

It is important to note that in the strictly comparative studies done here, all five proteins respond in the same manner to denaturant / viscogen. In all cases the position of the TS (relative to N) is unaffected by the viscogen, the denatured states show similar evidence for collapse and the free energy of unfolding is affected in the same way (Supplementary Figs 2 & 4).

Kinematic viscosity was measured using U-tube viscometers (Poulten Selfe & Lee Ltd), and multiplied by the density to find the dynamic viscosity.

Methods

All experiments were carried out at 25 ± 0.1 °C in 50 mM sodium phosphate buffer (except the R16o15c Φ -value analysis carried out at 10°C). For R17 and R17o15c 5 mM DTT was added to the buffers.

Fitting kinetic data

R16 and R16o15c have a single proline residue, and so refolding data were described by a double exponential equation, with the major, fast folding phase accounting for ~ 80% of the amplitude. The slow phase has been shown to result from proline isomerisation⁷. All unfolding data and refolding data for R15, R17, and R17o15c were well described by a single exponential process.

All the chevron plots, with fits, are shown in Supplementary Fig. 3. For each domain, the sets of eight chevron plots at varying [glucose] were fitted globally. This was both to reduce error in the fits, and to allow fitting of the curvature seen in all the chevrons except those of R15.

Data collection for R15 was limited by the fast folding and the dead time of our stopped-flow instrument. As a result, the arms of the R15 chevron plot are short (Supplementary Fig. 3a), and the curvature we have previously inferred in the unfolding arm is not seen¹⁰, so the chevrons were fitted to a linear chevron fit. The very short arms made accurate fitting of the gradient of the folding arm, m_{kf} , and the gradient of the unfolding arm, m_{ku} , difficult. Individual fitting showed that m_{ku} is unaffected by increasing [glucose], but m_{kf} decreases (Supplementary Fig. 4a). This leads to a decrease in the kinetic m -value, m_{kin} , where $m_{kin} = RT(m_{kf} + m_{ku})$. This decrease is comparable to the decrease in equilibrium m -value, m_{eq} , with increasing [glucose] seen in the equilibrium data (Supplementary Fig. 4a). As $m_{kin} = m_{eq}$, within error, for R15, m_{kf} was shared, and m_{ku} constrained such that $RT(m_{kf} + m_{ku}) = m_{eq}$ in the global fitting of R15 (Supplementary Fig. 3a; Supplementary Fig. 4b).

Because R16 and R17 fold so much slower than R15, curvature can clearly be seen in the unfolding chevron arms. These data are best fitted using a sequential transition state model³¹⁻³⁴. R16 has been fitted in this manner previously^{8,35} and the longer unfolding arms seen for R17 through the use of the denaturant GdmCl, rather than urea, makes it possible to fit R17 in the same way. The fitting was carried out as described in⁸. As in R15 the unfolding m -values (in this case m_{-1} , m_2 and m_{-2}) were shared, for each set of eight chevron plots (Supplementary Fig. 3b,c). The m -values shared correspond to the m_{ku} that was shared for R15. In R16 the kinetic and equilibrium m -values were similar which is an indication that the fitting method is appropriate. In R17, although $m_{kin} > m_{eq}$ (Supplementary Fig. 4d), this has been seen before for R17 using other fitting methods⁹, and is a characteristic of the domain, so was not considered further here.

Curvature was seen in both chevron arms for R16o15c. At high [glucose] the curvature in the refolding arm became negative and is probably due to aggregation of a non-evolved domain. This curvature has been excluded from the analysis. Although R16o15c folds faster than R16 and R17, enough curvature was seen in the unfolding arm to allow fitting using the sequential transition state model used above, again sharing the unfolding m -values. For R17o15c, short arms displayed little or no curvature, and longer arms displayed significant curvature, irrespective of which arm they are. Because of this a broad transition state barrier model was considered the most appropriate fitting method to use³⁵⁻⁴⁰. This describes the rate limiting transition state moving towards the native state as the concentration of denaturant increases. In this fitting method a second-order polynomial term is added to a two-state chevron fit to account for the curvature seen in the chevron limbs. The eight chevrons were fitted globally, and the curvature term and m_{ku} shared globally. Sharing these two terms was not necessary for the fitting, but reduced fitting errors considerably.

The isostability approach to viscosity dependence experiments

We note that the commonly used isostability approach^{16-18,20,21}, using denaturant to counteract the stabilising effects of the viscogen, has been questioned on the grounds that the two may not be directly additive⁴¹. It is therefore most important to note that here, in these strictly comparative studies, all five proteins respond in the same manner to denaturant / viscogen. In all cases the position of the TS (relative to N) is unaffected by the viscogen, the denatured states show similar evidence for collapse and the free energy of unfolding is affected in the same way

Core-swap design, production and characterisation

The pdb structure of R15, R16 and R17 as a tandem repeat, 1u4q⁴², was used to determine which residues had side chains with 15% solvent accessible surface area (SASA) which were defined as core. By comparing across the three domains, outliers due to very large / small side chains were excluded, and the same 35 residues were defined as core for all three

domains (Supplementary Fig. 5; Fig. 1b,c). The core residues that were not identical between the domain pairs were identified and synthetic genes produced. R16o15c (a protein with the outside of R16 and the core of R15) and R15o16c were made using overlapping primers and standard PCR techniques, and R17o15c and R15o17c were purchased from GenScript. Each was inserted into the modified pRSETA vector used for R15, R16 and R17, and expressed, purified and characterised (for R16o15c and R17o15c) as for its major parent. R15o16c and R15o17c were expressed insolubly and could not be easily refolded.

Φ-value analysis of R16o15c

Φ-value analysis is a powerful method for investigating the structure of the transition state (TS) for folding⁴³. The regions of the protein that are significantly structured in the TS will have high Φ-values, and the Φ-values in unstructured regions are low. The Φ-value analysis of R16o15c was carried out in a very similar manner to those of R15 and R16^{8,10}. The structure of R16o15c has not been determined, however the similarity to its parents in terms of purification and biophysical characteristics indicates that the structures are similar. Consequently, the same positions, and types of mutations made, were chosen as the two parents. Furthermore, the change in stability on mutation for the core residues, $\Delta\Delta G_{D-N}$, was the same as observed in R15, suggesting that the core structure was maintained (Supplementary Fig. 8). All experimental work was carried out as for R15, including working at 10 °C, to access longer chevron arms than is possible at 25 °C. Urea was used as the denaturant, to allow direct comparison with the Φ-value analysis of the two parent proteins. Equilibrium curves were analysed as described⁴⁴ and kinetic traces were analysed as described for R16^(above and 10). The fitting of the chevron plots was carried out as describe for wild-type in GdmCl (above) using the sequential transition state model and sharing the unfolding *m*-values (Supplementary Fig. 7). Folding Φ-values, Φ_f^{2M} , were calculated at 2 M urea, to avoid long extrapolations (Supplementary Table 2; Fig. 4a):

$$\Phi_f^{2M} = \frac{RT \ln \left(\frac{k_{f,WT}^{2M}}{k_{f,mut}^{2M}} \right)}{\Delta\Delta G_{D-N}} \quad (4)$$

where $k_{f,WT}^{2M}$ and $k_{f,mut}^{2M}$ are the folding rate constants at 2 M urea for wild-type and mutant proteins, respectively. For surface Ala-Gly Φ-values Ala was used as the reference (wild-type) and Gly the mutant.

The Φ-values were calculated for the first TS, which is the rate limiting TS for WT and all mutants both in water and at 2 M urea. This TS was compared with the equivalent TS in R16, and the rate limiting TS in water for R15 (Fig. 4). Note that R17o15c is not amenable to Φ-value analysis because it is too unstable to tolerate large deletion mutations.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by the Wellcome Trust (Grant number 064417/Z/01/A). B.G.W. was supported by an MRC studentship. J.C. is a Senior Wellcome Trust Research Fellow. We thank William Eaton, Peter Wolynes, Robert Best and Ben Schuler for very helpful discussions.

References

1. Bryngleson JD, Wolynes PG. Intermediates and barrier crossing in a random energy model (with applications to protein folding). *J. Phys. Chem.* 1989; 93:6902–6915.
2. Gromiha MM, Selvaraj S. Comparison between long-range interactions and contact order in determining the folding rate of two-state proteins: application of long-range order to folding rate prediction. *J. Mol. Biol.* 2001; 310:27–32. [PubMed: 11419934]
3. Plaxco KW, Simons KT, Baker D. Contact order, transition state placement and the refolding rates of single domain proteins. *J. Mol. Biol.* 1998; 277:985–994. [PubMed: 9545386]
4. Henry ER, Eaton WA. Combinatorial modeling of protein folding kinetics: Free energy profiles and rates. *Chem. Phys.* 2004; 307:163–185.
5. Munoz V, Eaton WA. A simple model for calculating the kinetics of protein folding from three-dimensional structures. *Proc. Natl Acad. Sci. USA.* 1999; 96:11311–11316. [PubMed: 10500173]
6. Clarke J, Cota E, Fowler SB, Hamill SJ. Folding studies of immunoglobulin-like beta-sandwich proteins suggest that they share a common folding pathway. *Structure.* 1999; 7:1145–1153. [PubMed: 10508783]
7. Scott KA, Batey S, Hooton KA, Clarke J. The folding of spectrin domains I: wild-type domains have the same stability but very different kinetic properties. *J. Mol. Biol.* 2004; 344:195–205. [PubMed: 15504411]
8. Scott KA, Randles LG, Clarke J. The folding of spectrin domains II: phi-value analysis of R16. *J. Mol. Biol.* 2004; 344:207–221. [PubMed: 15504412]
9. Scott KA, Randles LG, Moran SJ, Daggett V, Clarke J. The folding pathway of spectrin R17 from experiment and simulation: using experimentally validated MD simulations to characterize states hinted at by experiment. *J. Mol. Biol.* 2006; 359:159–173. [PubMed: 16618492]
10. Wensley BG, Gärtner M, Choo W, Batey S, Clarke J. Different members of a simple three-helix bundle protein family have very different folding rate constants and fold by different mechanisms. *J. Mol. Biol.* 2009; 390:1074–1085. [PubMed: 19445951]
11. Kramers HA. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica.* 1940; 7:284–304.
12. Sutto L, Lätzer J, Hegler JA, Ferreira DU, Wolynes PG. Consequences of localized frustration for the folding mechanism of the IM7 protein. *Proc. Natl Acad. Sci. USA.* 2007; 104:19825–19830. [PubMed: 18077415]
13. Cellmer T, Henry ER, Hofrichter J, Eaton WA. Measuring internal friction of an ultrafast-folding protein. *Proc. Natl Acad. Sci. USA.* 2008; 105:18320–18325. [PubMed: 19020085]
14. Jas GS, Eaton WA, Hofrichter J. Effect of viscosity on the kinetics of α -helix and β -hairpin formation. *J. Phys. Chem. B.* 2001; 105:261–272.
15. Ansari A, Jones CM, Henry ER, Hofrichter J, Eaton WA. The role of solvent viscosity in the dynamics of protein conformational changes. *Science.* 1992; 256:1796–1798. [PubMed: 1615323]
16. Chrnyk BA, Mathews CR. Role of diffusion in the folding of the alpha subunit of tryptophan synthase from *Escherichia coli*. *Biochemistry.* 1990; 29:2149–2154. [PubMed: 2183877]
17. Jacob M, Schindler T, Balbach J, Schmid FX. Diffusion control in an elementary protein folding reaction. *Proc. Natl Acad. Sci. USA.* 1997; 94:5622–5627. [PubMed: 9159122]
18. Plaxco KW, Baker D. Limited internal friction in the rate-limiting step of a two-state protein folding reaction. *Proc. Natl Acad. Sci. USA.* 1998; 95:13591–13596. [PubMed: 9811844]
19. Frauenfelder H, Fenimore PW, Chen G, McMahon BH. Protein folding is slaved to solvent motions. *Proc. Natl Acad. Sci. USA.* 2006; 103:15469–15472. [PubMed: 17030792]
20. Pradeep L, Udgaonkar JB. Diffusional barrier in the unfolding of a small protein. *J. Mol. Biol.* 2007; 366:1016–1028. [PubMed: 17188296]
21. Qiu LL, Hagen SJ. Internal friction in the ultrafast folding of the tryptophan cage. *Chem. Phys.* 2004; 312:327–333.
22. Pabit SA, Roder H, Hagen SJ. Internal friction controls the speed of protein folding from a compact configuration. *Biochemistry.* 2004; 43:12532–12538. [PubMed: 15449942]

23. Zwanzig R. Diffusion in a rough potential. *Proc. Natl Acad. Sci. USA.* 1988; 85:2029–2030. [PubMed: 3353365]
24. Godoy-Ruiz R, Henry ER, Kubelka J, Hofrichter J, Munoz V, Sanchez-Ruiz JM, Eaton WA. Estimating free energy barrier heights for an ultrafast folding protein from calorimetric and kinetic data. *J. Phys. Chem. B.* 2008; 112:5938–5949. [PubMed: 18278894]
25. Nettels D, Gopich IV, Hoffmann A, Schuler B. Ultrafast dynamics of protein collapse from single-molecule photon statistics. *Proc. Natl Acad. Sci. USA.* 2007; 104:2655–2660. [PubMed: 17301233]
26. Yang WY, Gruebele M. Folding at the speed limit. *Nature.* 2003; 423:193–197. [PubMed: 12736690]
27. Gruebele, M. Protein folding misfolding and aggregation: Classical themes and novel approaches. Munoz, V., editor. Royal Society of Chemistry; 2008. p. 106-138.
28. Pogorelov TV, Luthey-Schulten Z. Variations in the fast folding rates of the lambda-repressor: A hybrid molecular dynamics study. *Biophys. J.* 2004; 87:207–214. [PubMed: 15240458]
29. 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; 100:13286–13291. [PubMed: 14595026]
30. Johnson CP, Gaetani M, Ortiz V, Bhasin N, Harper S, Gallagher PG, Speicher DW, Discher DE. Pathogenic proline mutation in the linker between spectrin repeats: disease caused by spectrin unfolding. *Blood.* 2007; 190:3538–3543. [PubMed: 17192394]

Additional References

31. Bachmann A, Kiefhaber T. Apparent two-state tandem folding is a sequential process along a defined route. *J. Mol. Biol.* 2001; 306:375–386. [PubMed: 11237606]
32. Kiefhaber T, Kohler HH, Schmid FX. Kinetic coupling between protein folding and prolyl isomerization. I. Theoretical models. *J. Mol. Biol.* 1992; 224:217–229. [PubMed: 1548700]
33. Sanchez IE, Kiefhaber T. Evidence for sequential barriers and obligatory intermediates in apparent two-state protein folding. *J. Mol. Biol.* 2003; 325:367–376. [PubMed: 12488101]
34. Szabo, ZG. *Comprehensive Chemical Kinetics.* Bamford, CH.; Tipper, CFH., editors. Vol. 2. Elsevier; 1969. p. 1-81.
35. Scott KA, Clarke J. Spectrin R16: broad energy barrier or sequential transition states? *Protein Sci.* 2005; 14:1617–1629. [PubMed: 15930007]
36. Oliveberg M. Characterisation of the transition states for protein folding: towards a new level of mechanistic detail in protein engineering analysis. *Curr. Opin. Struct. Biol.* 2001; 11:94–100. [PubMed: 11179897]
37. Oliveberg M, Tan YJ, Silow M, Fersht AR. The changing nature of the protein folding transition state: implications for the shape of the free-energy profile for folding. *J. Mol. Biol.* 1998; 277:933–943. [PubMed: 9545382]
38. 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; 38:6499–6511. [PubMed: 10350468]
39. Silow M, Oliveberg M. High-energy channeling in protein folding. *Biochemistry.* 1997; 36:7633–7637. [PubMed: 9201903]
40. Ternstrom T, Mayor U, Akke M, Oliveberg M. From snapshot to movie: phi analysis of protein folding transition states taken one step further. *Proc. Natl Acad. Sci. USA.* 1999; 96:14854–14859. [PubMed: 10611302]
41. Ladurner AG, Fersht AR. Upper limit of the time scale for diffusion and chain collapse in chymotrypsin inhibitor 2. *Nat. Struct. Biol.* 1999; 6:28–31. [PubMed: 9886288]
42. Kusunoki H, Minasov G, Macdonald RI, Mondragon A. Independent movement, dimerization and stability of tandem repeats of chicken brain alpha-spectrin. *J. Mol. Biol.* 2004; 344:495–511. [PubMed: 15522301]

43. Fersht AR, Matouschek A, Serrano L. The folding of an enzyme. I. Theory of protein engineering analysis of stability and pathway of protein folding. *J. Mol. Biol.* 1992; 224:771–782. [PubMed: 1569556]
44. Pace CN. Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods Enzymol.* 1986; 131:266–280. [PubMed: 3773761]

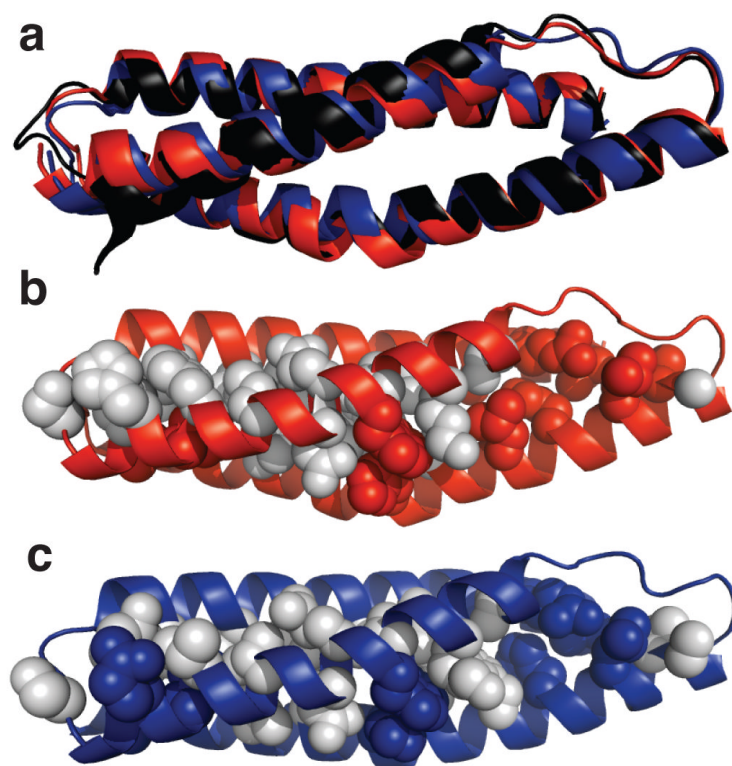


Figure 1. Spectrin domains used in this study. (a) Superposition of R15 (black), R16 (red) and R17 (blue) (pairwise backbone RMSD <math>< 1 \text{ \AA}</math>). (b) R16o15c. Core residues (15% SASA) shown in space-filling mode. Red, residues identical in R16 and R15; grey, the 22 residues that were substituted into R16 from R15. Sequence identity of R16o15c to “major parent”, R16, 81% and to “minor parent”, R15, 51%. (c) R17o15c. Blue, residues identical in R17 and R15; grey, the 24 residues that were substituted into R17 from R15. Sequence identity of R17o15c to major parent, R17, 79% and to minor parent, R15, 53%.

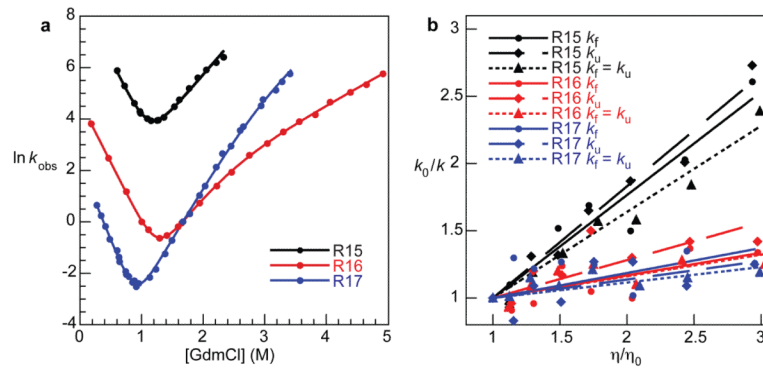


Figure 2.

Kinetic data for R15, R16 and R17 spectrin domains. (a) Chevron plots in 0 M glucose. R15 folds and unfolds significantly faster than R16 and R17. (b) Dependence of the relative rate constants (k_0/k) on relative solvent viscosity (η/η_0) (subscript 0 indicates rate constant / viscosity in 0 M glucose). k_f and k_u determined at $\Delta G = 1.5$ kcal/mol, and $k_f = k_u$ at $\Delta G = 0$ kcal/mol. R16 and R17 rate constants show significantly lower viscosity dependence.

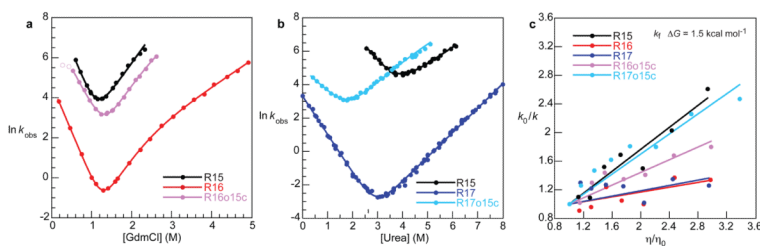


Figure 3.

Kinetic data for core-swapped proteins. Core-swapped proteins coloured according to major parent, R16o15c, pink (major parent R16), R17o15c, light blue (major parent R17). (a) and (b) Chevron plots compared to parent proteins. Both core-swapped proteins fold and unfold significantly faster than major parent, even though R17o15c is significantly destabilised. (c) Dependence of the relative rate constants for folding at $\Delta G = 1.5$ kcal/mol (k_f/k) on relative solvent viscosity (η/η_0). Both core swapped proteins have a solvent viscosity dependence significantly greater than the major, slow-folding parent. See also Supplementary Fig. 6.

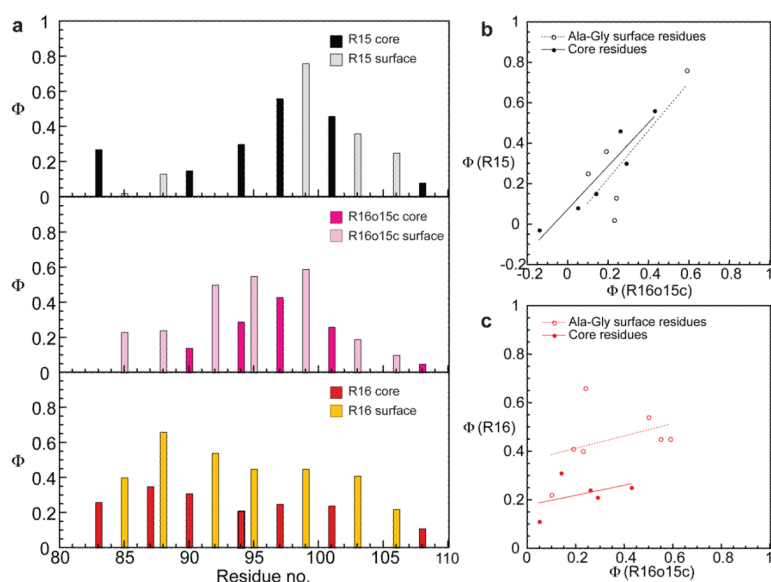


Figure 4.

Φ -values of R16o15c resemble R15. A Φ -value is a measure of the amount of structure formed around a specific residue in the TS. High Φ -values indicate significant structure, low Φ -values little or no structure. (a) Histograms of Φ -values for the C-helix of R15 (top) R16o15c (centre) and R16 (bottom). The dark bars represent truncation of core residues to Ala, to probe formation of tertiary structure; the light bars are surface Ala to Gly mutations to probe helix formation. In R16 the Φ -values monitoring helix formation are significantly higher throughout the helix than the core Φ -values. This is good evidence to suggest that helix formation precedes helix packing. (b) There is a good correlation between the Φ -values in R15 and R16o15c for both core ($R = 0.94$) and surface ($R = 0.78$) residues. (c) There is no significant correlation between the Φ -values in R16 and R16o15c for either core ($R = 0.41$) or surface ($R = 0.38$) residues.