doi:10.1006/jmbi.2000.4352 available online at http://www.idealibrary.com on IDEAL® J. Mol. Biol. (2001) 306, 555–563

JMB

Circularization Changes the Folding Transition State of the src SH3 Domain

Viara P. Grantcharova and David Baker*

Department of Biochemistry
University of Washington
Seattle, WA 98195, USA

Native state topology has been implicated as a major determinant of protein-folding mechanisms. Here, we test experimentally the robustness of the src SH3-domain folding transition state to changes in topology by covalently constraining regions of the protein with disulfide crosslinks and then performing kinetic analysis on point mutations in the context of these modified proteins. Circularization (crosslinking the N and C termini) of the src SH3 domain makes the protein topologically symmetric and causes delocalization of structure in the transition state ensemble suggesting a change in the folding mechanism. In contrast, crosslinking a single structural element (the distal β-hairpin) which is an essential part of the transition state, results in a protein that folds 30 times faster, but does not change the distribution of structure in the transition state. As the transition states of distantly related SH3 domains were previously found to be very similar, we conclude that the free energy landscape of this protein family contains deep features which are relatively insensitive to sequence variations but can be altered by changes in topology.

Keywords: protein folding; folding kinetics; folding mechanism; transition state; SH3 domain

*Corresponding author

Introduction

A recent development in the protein folding field has been the empirical observation that native state topology is a major determinant of folding rates, with simple fold proteins folding faster than proteins with complicated topologies.[1] The remarkable correlation found between the average sequence separation of interacting residues in the native structure (contact order) and the rate of folding suggests that the free energy barrier to folding has a large entropic contribution while variations in the strength of the stabilizing interactions manifest themselves on a smaller scale. Consistent with the idea that the molecular details of the interactions are overshadowed by the entropic cost of making them, several theoretical models have been successful in predicting the transition state for folding and/or the folding rate for small proteins using only information from the native state structure.[2–5] Furthermore, recent experimental studies have established the conservation of folding transition states among homologous proteins with the same topology but sequence identity as low as 13%[6–8] suggesting that once the native-state topology is specified by the sequence the folding transition state is largely determined as well. The goal of our study is to explore and test these conclusions further.

Previous studies of the folding transition state of the src SH3 domain showed that it involves the association of the distal β-hairpin and the diverging turn, while the N and C termini are completely disordered (Figure 1(a)).[8,9] Here, we explore the robustness of this transition state to chain crosslinks in order to test the role of topology. Previously Serrano and co-workers showed that circularly permuting the α-spectrin SH3 domain can change its transition state depending on the site of permutation,[10] while smaller mutations that stabilize a part of the folding nucleus do not alter structure elsewhere in the transition state.[7,11] This argued for a conformationally restricted transition state, which requires the interaction of specific

Present address: V. P. Grantcharova, Center for Genomics Research, Harvard University, 16 Divinity Ave, Cambridge, MA 02138, USA.

Abbreviations used: SH3, src homology 3; Gnd, guanidine; wt, wild-type; CO, contact order; NC, circular src SH3 domain; SS, src SH3 mutant with a distal hairpin crosslink; CI2, chymotrypsin inhibitor 2.

E-mail address of the corresponding author: dabaker@u.washington.edu

0022-2836/01/030555–9 $35.00/0

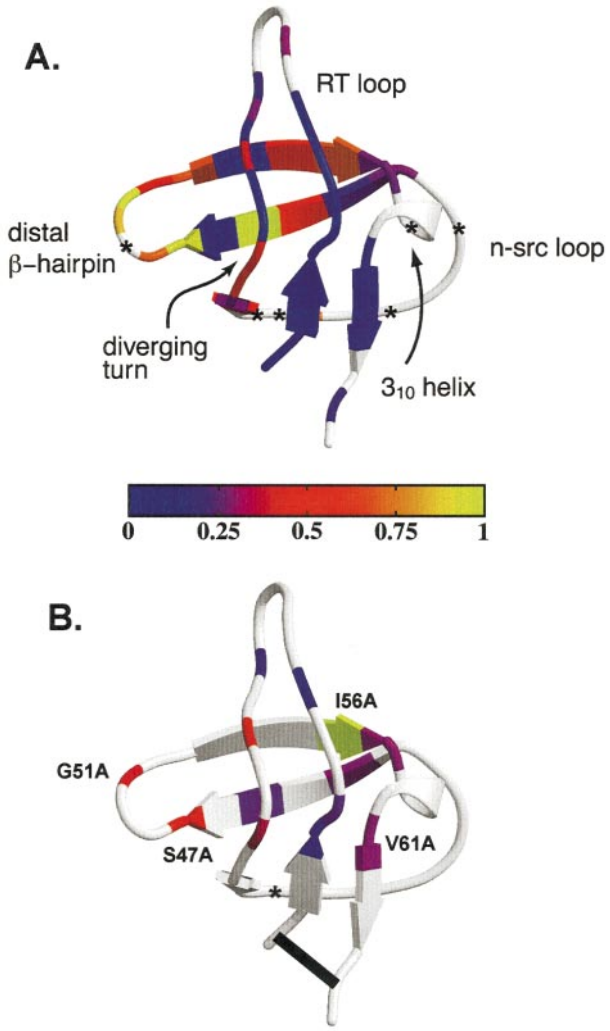

Figure 1. Structure in the transition state of (a) wt, (b) NC terminal crosslink. Color scheme is continuous from yellow ($\Phi_F = 1$) to red ($\Phi_F = 0.5$) to blue ($\Phi_F = 0$). Asterisks indicate negative Φ_F values, which suggest the involvement of these residues in non-native interactions in the transition state. Graphics were generated with Molscript[28] and Raster3d.[29,30]

parts of the molecule to overcome the loss of entropy during folding. Here, we examine the effects on the transition state of disulfide crosslinking the distal β-hairpin and circularizing (linking the N and C termini) the protein. These modifications were previously characterized kinetically in their reduced and oxidized forms to test backbone conformational ordering in the transition state.[12] Crosslinking the distal hairpin increased the folding rate 30-fold without affecting the unfolding rate, suggesting that this structural element is as conformationally constrained in the transition state as in the native state. Crosslinking the N and the C termini stabilized the protein significantly both by increasing the folding rate and by decreasing the unfolding rate, indicating that the two termini are not fully interacting in the transition state. Here, we perform mutational analysis to determine if these modifications affect the distribution of structure in the transition state. By deleting parts of individual residues (as in mutations to an alanine residue) and then assessing the effect of the mutation on stability and folding kinetics, we can gain site-specific information about structure at the rate-limiting step.[13] The degree of structure formation around each residue in the transition state can be conveniently represented by ϕ_F values, defined as $\Delta\Delta G_{U\text{-}\ddagger}/\Delta\Delta G_{U\text{-}F}$ (see Materials and Methods). In the case of the distal hairpin crosslink, we investigate whether it allows overcoming of the entropic cost of ordering earlier in the folding reaction and thus makes other parts of the transition state less structured. In the case of the terminal crosslink, we examine if significantly decreasing the entropic barrier to folding and making the protein topologically symmetric causes its transition state to become delocalized with all residues contributing equally, or whether it remains structurally polarized. A similar circularization experiment was performed on chymotrypsin inhibitor 2 (CI2),[14] however, circularization did not affect its transition state probably because it is largely delocalized even in the wild-type protein.[15] Our results suggest that, at least for the src SH3 domain, the transition state ensemble can be shifted when the native topology is significantly perturbed as in circularization, but not by stabilization of the existing nucleus.

Results

Disulfide crosslinking of the distal β-hairpin

Covalent crosslinking of the distal β-hairpin is expected to decrease the entropy of the denatured state and stabilize intrahairpin interactions. If both chain entropy and energy are smoothly varying functions of the degree of ordering and the position of the transition state is determined by their imperfect cancellation, then such a change should alter the position of the transition state. Consistent with the Hammond postulate,[16] destabilizing the denatured state would shift the position of the transition state closer to the denatured state and result in lower Φ_F values in regions other than the distal hairpin. In contrast, if the energy decreases abruptly when a large number of contacts form simultaneously, the transition state would be less sensitive to changes in interaction strengths and it would be effectively ''locked''. In that case, we would expect that crosslinking the distal hairpin will increase the folding rate, but structure in the transition state will remain the same. The distal hairpin was previously crosslinked by mutating residues W43 and S58 to cysteine residues and forming a disulfide bridge between them under oxidizing conditions.[12] Here, we perform Φ value analysis on several mutants throughout the crosslinked protein (denoted SS) to determine whether

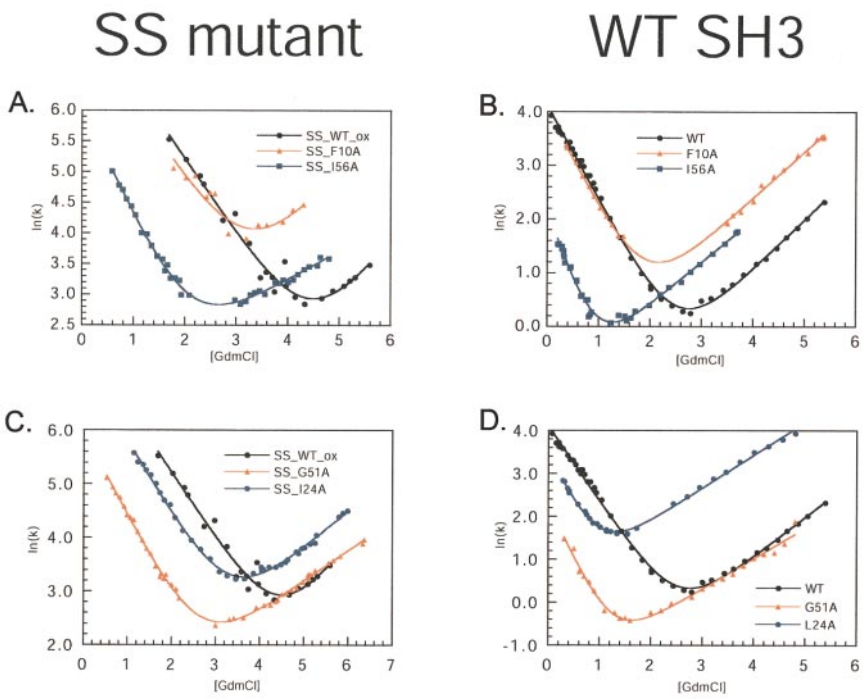

Figure 2. Kinetic analysis of mutants in the context of the distal hairpin SS crosslink. Rates of folding and unfolding were measured using stopped flow fluorescence at 295 K. Continuous lines represent the best fit to the experimental data (Kaleidagraph).

structure in the transition state has changed (F10A and A12G (N-terminal strand), D15A and L24A (RT loop), G29A and E30A (diverging turn), I34A (n-src loop) and G51A and I56A (distal β-hairpin)). The relative effect of the mutations on the rate of folding and unfolding is very similar in the context of the crosslinked protein and in the wild-type (Figure 2 and Table 1) resulting in similar Φ_F values. Even though the distal hairpin is covalently crosslinked at its base, G51A retains a Φ_F value close to 1, suggesting that this residue plays an important role in organizing structure at the turn (the glycine may allow the sidechain oxygens of S47 and T50 to hydrogen bond with neighboring backbone amides, positioning them to interact with the diverging turn). The only mutations that exhibit different behavior in the wt and crosslinked protein are L24A and I34A, but these effects can be attributed to changes in local structure around the disulfide crosslink. In the wild-type (wt) protein, I34A destabilizes the transition state more than the native state, suggesting non-native structure in the transition state, however, in the crosslinked mutant the I34A mutation has an intermediate ϕ_F value, perhaps because the replacement of the large W43 sidechain by cysteine destabilizes non-native structure in this region. L24A, on the other hand, shows an increase in Φ_F value in the crosslinked mutant from 0.26 to 0.42 indicating that the diverging turn is slightly better packed onto the distal hairpin in the transition state. Overall, however, the transition state of the crosslinked protein involves the same

Table 1. Kinetic parameters for the SS mutants

Mutant	$\ln(k_f)^{1.5\ M}$	$\ln(k_u)^{5\ M}$	m_f	m_u	$\Delta\Delta G_U$	Φ_F^{SS}	Φ_F^{WT} [b]
SS_WT[a]	5.83	2.83	0.716	0.600	-	-	-
SS_F10A	5.50	4.98	0.723	0.519	−1.45	0.13	0.10
SS_D15A	5.70	3.90	0.666	0.358	−0.721	0.13	−0.22
SS_L24A	5.12	3.80	0.803	0.404	−0.985	0.42	0.26
SS_G29A	4.42	5.82	0.630	0.638	−2.58	0.32	0.44
SS_E30A	4.10	4.35	0.581	0.482	−1.91	0.53	0.62
SS_G51A	3.64	3.19	0.911	0.332	−1.50	0.86	1.06
SS_I56A	3.36	3.71	1.06	0.276	−1.96	0.74	0.71

k_f is reported in 1 M guanidine, while k_u is in 6 M guanidine to avoid extrapolation; m_f and m_u are the dependences of the folding and the unfolding rates, respectively, on Gnd. Typical errors for the kinetic measurements are 2-20 % as reported by Riddle *et al.*[8]

[a] Kinetic data for this mutant was published previously by Grantcharova *et al.*[12]

[b] Φ_F values taken from the paper by Riddle *et al.*[8]

structural elements as that of the wild-type SH3 domain. We can conclude that even though formation of the distal β-hairpin is required for the overcoming of the activation barrier, it is not sufficient, even when it is largely stabilized. The rate-limiting step involves bringing the distal β-hairpin and the diverging turn together to form a three-stranded β-sheet. Thus, stabilization of this element speeds folding, but does not alter the transition state ensemble.

Disulfide crosslinking of the N and C termini

Theoretical models of the transition state for folding emphasize the balance between loss of configurational entropy and formation of stabilizing interactions in determining which part of the molecule folds first.[17–19] The combination of structural elements in the protein that can bury the most surface area while losing the least amount of configurational entropy may nucleate folding. In the modeled free energy landscape for the src SH3 domain[2] there is only one set of segments (the distal hairpin and the diverging turn) which can associate with sufficient number of favorable contacts to compensate for the loss in entropy; all other pairings are entropically too costly and poorly populated to lead to productive folding. In particular, the two terminal strands, which form a sheet in the native state, were found to be completely unstructured in the transition state due to their large sequence separation. Our strategy here is to connect the termini and examine how the distribution of structure in the transition state changes. A circularized version of the src SH3 domain (denoted NC protein) was previously constructed by mutating both residues T9 and S64 to cysteine residues and forming a disulfide bridge between them under oxidizing conditions.[12] Crosslinking makes the topology of the protein symmetric and entropically there is no reason why one three-stranded sheet would form first over the other. One prediction is that circularization would greatly reduce structural polarization because it will offer alternative routes for folding. Another possibility is that the same folding nucleus will be maintained because the interactions present in it are inherently more favorable. Such a breakdown in symmetry is seen in protein L which is topologically symmetric and yet one part of the molecule is preferentially structured at the transition state.[20,21] Distinguishing between these two possibilities addresses the relative importance of variations in interaction energies and chain entropy in determining the folding transition state.

In order to determine the effect of circularization on the transition state we examined the effect of mutants in the context of the crosslinked protein. A total of 14 mutants were designed to probe different regions of the transition state: A12G, L13A, Y16A and D23A (first strand) and RT loop; F26A and G29A (diverging turn); I34A and W43A (n-src loop); A45G, S47A and G51A (distal hairpin); I56A and P57A (3_{10} helix); V61A (C-terminal strand). Kinetic analysis (Figure 3 and Table 2) reveals that there are clear differences between some of the mutants in the NC protein and the corresponding mutations in the wt.[8] The most significant changes are observed in the Φ_F values of residues in the distal hairpin. A45G, S47A and G51A affect both the rates of folding and unfolding in the NC mutant, while in the context of the wt protein their Φ_F values were all 1 (Figure 3(a) and (b)). Since we are not certain of the homogeneity of

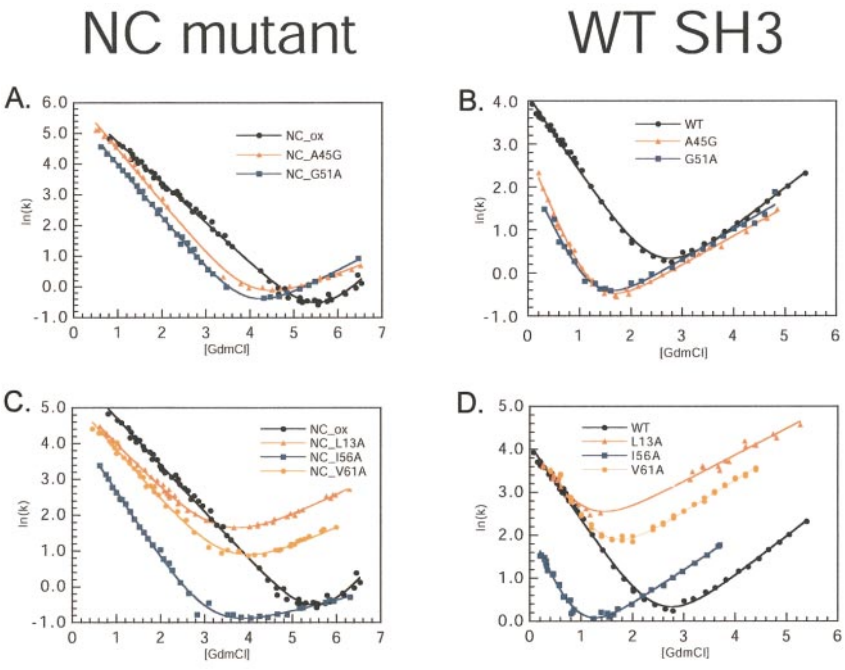

Figure 3. Kinetic analysis of mutants in the context of the NC terminal crosslink. Rates of folding and unfolding were measured using stopped flow fluorescence at 295 K. Continuous lines represent the best fit to the experimental data (Kaleidagraph).

Table 2. Kinetic parameters for the NC crosslink mutants

Mutant	$\ln(k_f)^{1\ M}$	$\ln(k_u)^{6\ M}$	m_f	m_u	$\Delta\Delta G_U$	Φ_F^{NC}	Φ_F^{WT} [b]
NC_WT[a]	4.73	−0.597	0.766	0.800	-	-	-
NC_A12G	4.41	2.76	0.874	0.637	−2.16	0.09	0.05
NC_L13A	4.00	2.58	0.740	0.340	−2.29	0.19	−0.03
NC_Y16A	4.01	3.28	0.752	0.431	−2.69	0.16	0.03
NC_D23A	4.64	0.305	0.760	0.500	−0.580	0.09	0.13
NC_F26A	3.26	1.13	0.812	0.370	−1.87	0.46	0.40
NC_G29A	3.46	1.78	0.780	0.500	−2.14	0.35	0.44
NC_I34A	3.66	−1.22	1.082	0.216	−0.263	[c]	[c]
NC_W43A	3.97	1.75	0.830	0.337	−1.82	0.25	0.15
NC_A45G	4.50	0.436	1.00	0.322	−0.737	0.18	1.20
NC_S47A	3.95	0.400	0.990	0.309	−1.04	0.44	0.95
NC_G51A	3.99	0.551	0.994	0.437	−1.11	0.39	1.06
NC_I56A	2.67	−0.353	1.10	0.190	−1.35	0.89	0.71
NC_P57A	3.78	1.95	0.770	0.550	−2.05	0.27	0.24
NC_V61A	3.84	1.66	0.821	0.315	−1.84	0.28	−0.06

k_f is reported in 1 M guanidine, while k_u is in 6 M guanidine to avoid extrapolation; m_f and m_u are the dependences of the folding and the unfolding rates, respectively, on Gnd. Typical errors for the kinetic measurements are 2-20 % as reported by Riddle *et al.*[8]

[a] Kinetic data for this mutant was published by Grantcharova *et al.*[12]

[b] Φ_F values taken from the paper by Riddle *et al.*[8]

[c] Mutation decreases both k_f and k_u.

the transition state the intermediate Φ_F values can either mean that the interactions in which these residues participate are not completely formed in the transition state, or that the transition state ensemble consists of some conformations in which the distal hairpin is formed and some in which the hairpin is disordered. On the other hand, mutations which probe formation of the hairpin newly created by the crosslink and the region around the 3_{10} helix have increased Φ_F values, suggesting that these residues now contribute to stabilization of the transition state. I56, an integral part of the hydrophobic core exhibits an increase in Φ_F value from 0.7 to 0.89 (Figure 3(c) and (d)). In the n-src loop, mutation of the buried W43 to alanine had no effect on the rate of folding in the wt context, but in the NC protein it has a Φ_F value of 0.25. In a similar way, V61 (C-terminal strand), which takes part in the hydrophobic core, and L13 (N-terminal strand), which interacts with the C-terminal strand on the solvent exposed side, both have increased Φ_F values upon mutation from 0 to 0.28 and 0.19, respectively. Other mutations in the first strand and the RT loop (A12G and D23A) exhibit roughly the same Φ_F values in the NC mutant as in the WT protein (Φ_F values close to 0), suggesting that despite the crosslink this region remains unstructured in the transition state. F26A and G29A in the diverging turn also preserve their intermediate Φ_F values in the NC mutant.

Taken together, these data suggest that the transition state of the circularized protein is significantly different from that of the wt protein (Figure 1(a) and (b); Figure 4(b)). The transition state of the WT SH3 domain is highly polarized with the distal hairpin and the diverging turn almost fully ordered in a three-stranded sheet, and the termini disordered. In contrast, the circularized protein appears to have a more delocalized transition state with a prevalence of intermediate Φ_F values in most of the structural elements probed. The distal hairpin, however, is still more ordered than other hairpins in the protein, probably because it has the highest density of intrahairpin interactions. It is interesting that one residue, I56 (central hydrophobic core residue), stands out with a Φ_F value close to 1 and therefore can be viewed as the nucleus around which structure consolidates. We can surmise that because of the circular topology of the protein, hairpin formation is not as important in the NC protein as it is in the wt. Instead, it appears that hydrophobic collapse, rather than local β-hairpin and sheet formation,

Table 3. Kinetic parameters for WT and mutants in 0.4 M sodium sulfate

Mutant	$\ln(k_f)^{0.5\ M}$	$\ln(k_u)^{5\ M}$	m_f	m_u	$\Delta\Delta G_U$	Φ_F^{sulf}	Φ_F^{WT} [b]
WT_sulf	4.31	0.867	0.695	0.805	-	-	-
F10I_sulf	4.22	3.65	0.702	0.580	−1.68	0.03	−0.05
L44A_sulf	2.87	2.66	1.24	0.474	−1.89	0.45	0.54
G51A_sulf	2.33	1.12	0.959	0.519	−1.31	0.89	1.06
I56A_sulf	2.06	1.70	1.14	0.468	−1.81	0.73	0.71

k_f is reported in 0.5 M guanidine, while k_u is in 5 M guanidine to avoid extrapolation; m_f and m_u are the dependences of the folding and the unfolding rates, respectively, on Gnd. Typical errors for the kinetic measurements are 2-20 % as reported by Riddle *et al.*[8]

[a] Φ_F values taken from the work by Riddle *et al.*[8]

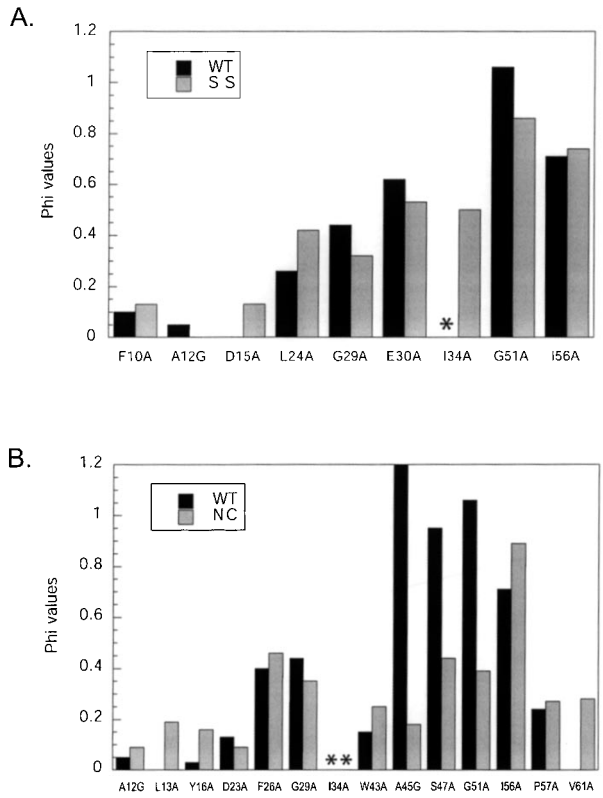

Figure 4. Comparison of Φ_F values for mutated residues in the wt and the (a) SS crosslink and (b) NC crosslink. Asterisk indicates a negative Φ_F value, which suggests the involvement of this residue in non-native interactions in the transition state.

drives the early stages of folding of the circularized protein, due to the decreased cost of bringing together residues distant in the chain.

Global stabilization and structure in the transition state

Probing the effect of a globally stabilizing agent on the rate-limiting step in folding provides another way to examine the robustness of the transition state. Sodium sulfate stabilizes proteins, presumably by its preferential hydration of water, therefore facilitating hydrophobic collapse.[22] Its effect on the kinetics of the src SH3 domain is to increase the folding rate and to decrease the unfolding rate (Table 3), indicating that protein desolvation occurs both before and after the transition state. It also decreases the denaturant dependence of the folding rate (i.e. m_f value), suggesting that it makes the denatured state more compact. We performed kinetic analysis in the presence of 0.4 M sodium sulfate of several mutants, in the wt context, which in the absence of sodium sulfate cover the full range of Φ_F values: F10I (Φ_F value of 0); G51A (Φ_F value of 1); L44A and I56A (intermediate Φ_F values). All mutants conserve their Φ values in the presence of sodium sulfate (Table 3), suggesting that the transition state ensemble has not been changed by addition of the salt. This is a further confirmation that the transition state for folding is determined by deep features of the SH3 energy landscape. Similar experiments with the α-spectrin SH3 domain show that variations in pH do not affect transition state structure.[7]

Discussion

Changing the structure of transition state ensembles

Point mutagenesis, which probes residue-specific interactions, and covalent modifications (glycine loop insertions and disulfide crosslinking), which test long-range order in the transition state revealed the conformationally restricted nature of the transition state ensemble of the src SH3 domain.[8,12] Here, we have explored in more detail the free-energy landscape of this protein and have determined how changes in chain configurational entropy and interaction strengths affect the distribution of structure in the transition state.

Covalent crosslinking is a convenient way of altering the entropic cost of contact formation. It reduces the average sequence separation between interacting residues (i.e. contact order (CO)) and is therefore expected to increase the rate of folding, as observed in the very good correlation between CO and rate of folding for all the characterized two state folding proteins.[23] Both the NC and SS-crosslinked mutants fall within the spread observed for natural proteins on the CO *versus* log(k) plot (data not shown). It should be noted, however, that the distal hairpin crosslink causes a larger increase in folding rate than the NC-terminal crosslink even though the contact order of the NC protein is smaller. This is because the folding rate is sensitive to chain entropy loss in the transition state rather than the native state. The distal loop crosslink reduces the entropy of the denatured state dramatically but has essentially no effect on that of the transition state (the loop is already formed), while the NC crosslink reduces the entropy of both the denatured state and the transition state. The overall correlation between contact order and folding rate suggests that on average the distribution of contacts (contact order) in the folding transition state follows that of the native structure, but in any specific case, the effect of a crosslink on the folding rate will depend on the transition state structure and not solely on the reduction in the native state contact order.

The finding that stabilization of local structure by distal hairpin crosslinking (Table 1, Figure 4(a)) and global stabilization by sodium sulfate (Table 3) do not alter the placement of the transition state along the reaction coordinate (as judged by Φ_F value distributions) indicates that there are some deep features in the energy landscape which are not altered by such changes. These results are con-

sistent with experiments on other SH3 domains. The distantly related src and α-spectrin SH3 domains exhibit very similar transition states,[7,8] and stabilizing mutations[11] and changes in pH[7] do not seem to affect transition state structure of the spectrin SH3 domain. It appears then, that SH3 domains allow quite large variations in sequence and experimental conditions with no change to the transition state probably because there are no alternative structural elements that can be sufficiently stabilized to become folding nuclei. On the other hand, modifying the topology of the protein, as in circularization, can significantly change the free energy landscape to favor alternative routes for folding. Similar conclusions were drawn from the circular permutation experiments on the α-spectrin SH3 domain.[10] Connecting the wt termini with a small peptide linker and introducing a cut in the distal hairpin resulted in a shift in the structure of the transition state towards the n-src loop and the hairpin formed by the old termini; the former distal hairpin was completely disordered at the rate-limiting step. (In contrast, circular permutations that did not involve the distal loop β-hairpin did not appear to change the folding transition state.) Therefore, shifts in transition state structure can occur when formerly distant elements are covalently linked to reduce the entropic cost of their interaction.

It should be noted that SH3 domains have more polarized folding transition states than other small proteins (CI2, ACBP, AcP, FKBP12). Therefore, changes in the structure of the folding transition state are more evident for the SH3 domains than for proteins with more delocalized folding transition state ensembles. A particularly well studied example of a protein with a more delocalized folding transition state is chymotrypsin inhibitor 2 (CI2), only one residue has a Φ_F value greater than 0.5. Drastic changes in the topology of CI2 through circular permutation or circularization[14] have relatively little effect on the folding transition state.

While topology plays an important role in determining a protein's folding mechanism, the distribution of interaction energies throughout the protein also affects structure in the transition state. Recent experiments demonstrate that the transition state conservation observed for sequence homologs of the SH3 domain does not hold for structural homologs. For example, drastic mutagenesis, which weakens the interaction energies throughout the protein can make the transition state delocalized. A sequence simplified mutant of the src SH3 domain made predominantly of five amino acid residues (I, K, E, A, G) was found to have a more delocalized transition state (distal hairpin is not fully formed), most likely because the interactions stabilizing the wt SH3 transition state are not strong enough in the simplified mutant to overcome the loss in entropy and residues from other parts of the protein have to participate (Q. Yi and D.B., unpublished results). Furthermore, the presence of destabilizing features in a particular structural element might be required for functional reasons. This results in a preferential switch in the structured parts of the transition state to other regions with more favorable interactions. PsaE, a structural homolog of the SH3 domain, has a large loop insertion at the distal hairpin (13 residues) required for its function in the photosynthetic center I of cyanobacteria.[24] The larger entropic cost of forming stabilizing interactions makes the transition state delocalized with high Φ_F values distributed throughout the protein (P. Bowers and D.B., unpublished results). Sso7d, a DNA binding protein from *Sulfolobus solfataricus*,[1] is another structural homolog of the SH3 domain. Its distal hairpin contains three glycine residues at the turn and two more in the β-strands required for function, and is not well ordered early in folding (R. Guerois and L. Serrano, personal communication). The n-src hairpin, on the other hand, is the most regular element of structure with a favorable hydrophobic/hydrophilic pattern along the strands and a canonical type I turn. The burial of hydrophobic surface area between the n-src loop and the C-terminal helix further favors these elements as a folding nucleus. Therefore, the topology of the SH3 fold appears to allow several alternative routes of folding.

Another example of a simple system in which the effects of topology and local structural propensity on the transition state have been examined is the GCN4-p1 coiled coil. Mutational analysis indicated that the folding of the dimeric coiled coil occurs *via* multiple pathways.[25] Variations in helical propensity along the helix can favor one pathway over the others (e.g. the C terminus of the GCN4 coiled coil has a higher helical propensity and has been shown to form early in folding).[26] Destabilizing one part of the helix (with A to G mutations) channels folding to the alternative pathways. However, crosslinking the two helices to form a monomer abolishes the symmetry, making it entropically more favorable for folding to start at the part of the helix proximal to the tether, even if it has the lowest helical propensity.[25] In the monomeric version of the coiled coil, the topological constraints on the chain effectively limit the number of folding pathways to one and make the transition state less sensitive to variations in secondary structure.

A similar dependence of the folding mechanism on the stability of individual structural elements is observed in two proteins with symmetrical topology: protein L and protein G (an α-helix packed against two β-hairpins forming a sheet). In the transition state of protein L the first hairpin packs against the α-helix, while in protein G the second hairpin is more structured.[20,21] The choice of hairpin appears to depend on the intrinsic stability of the hairpins. In protein L, the first hairpin has more favorable side-chain:main-chain hydrogen bonds, while the second hairpin contains three consecutive residues with positive ϕ angles. In protein G, on the other hand, the second hairpin has an

extensive hydrogen bond network. Using computational protein design methods, the order of events in the folding of protein L and protein G can be switched by selectively stabilizing the hairpin normally formed late in folding (S. Nauli, B. Kuhlman & D.B., unpublished results).

The ability to change the transition state for folding tests our understanding of the factors contributing to its formation and specificity. Our results with the circularized src SH3 domain and the experimental studies on other proteins highlight the interplay of topologic constraints and contact energy heterogeneity in determining the structure of the transition state ensemble.

Materials and Methods

Mutagenesis

Point mutagenesis was accomplished using the Quick Change Site-Directed mutagenesis kit (Stratagene, La Jolla, CA). Plasmids harboring the mutations were transformed into BL21 cells, and protein was overexpressed and purified.[9] The His·Tag® was not removed for the purpose of this study. All mutants were sequenced to ensure that the mutagenesis was successful and the purified proteins were analyzed by mass spectrometry to confirm that each mutation was the expected one.

Disulfide crosslinking

The design of the SS and NC crosslink mutants was described by Grantcharova *et al*.[12] For all the mutants disulfide bonds were oxidized in the presence of 20 mM potassium ferricyanide $K_3Fe(CN)_6$ for ten minutes at room temperature. Reactions were performed in the dark because $K_3Fe(CN)_6$ is light sensitive. Disulfide formation was confirmed using Ellman's reagent.

Biophysical analysis

Protein solutions (100 μM) were made in 50 mM sodium phosphate (pH 6). For the experiments in sodium sulfate, 0.4 M sulfate was added to the solutions. The kinetics of folding and unfolding were followed by tryptophan fluorescence on a Bio-Logic SFM-4 stopped-flow instrument at 295 K. The unfolding reaction for the wt protein was previously determined to behave as a two-state process,[27] and the kinetic and equilibrium data for the mutants were fit to a two-state model. For each mutant the free energy of folding is calculated as:

$$\Delta G_{U-F} = RT \ln(k_f/k_u)$$

where k_f and k_u are the rates of folding and unfolding, respectively, in the absence of denaturant. The difference in the free energy of folding ($\Delta\Delta G_{U-F}$) and in the folding activation energy ($\Delta\Delta G_{U-\ddagger}$) between the wt protein and each mutant are calculated as:

$$\Delta\Delta G_{U-F} = RT(\ln(k_f^{wt}/k_f^{mut}) + \ln(k_u^{mut}/k_u^{wt}))$$

and

$$\Delta\Delta G_{U-\ddagger} = RT \ln(k_f^{wt}/k_f^{mut})$$

where k_f and k_u are the rates of folding and unfolding, respectively, at denaturant concentrations experimentally accessible for that mutant. This method avoids the extrapolation of k_f and k_u to 0 M denaturant and therefore does not rely on the accurate determination of the m_f and m_u values (the denaturant dependence of k_f and k_u, respectively).

The parameter Φ_F is defined as:

$$\Phi_F = \Delta\Delta G_{U-\ddagger}/\Delta\Delta G_{U-F}$$

and is interpreted as the fraction of the mutated residue's interactions that are formed in the transition state. A Φ_F value of 1 indicates that all of a residue's interactions are formed in the transition state, whereas a Φ_F of 0 means that the residue does not make stabilizing interactions in the transition state.[13]

Acknowledgments

We thank Raphael Guerois and Luis Serrano for communicating to us their unpublished results on the Sso7d protein. We are grateful to members of the Baker group for their useful comments on the manuscript. This work was supported by a grant from the NIH and Young Investigator awards to D. B. from the NSF and the Packard Foundation.

References

1. Baumann, H., Knapp, S., Lundback, T., Ladenstein, R. & Hard, T. (1994). Solution structure and DNA-binding properties of a thermostable protein from the archaeon Sulfolobus solfataricus. *Nature Struct. Biol.* **1**, 808-819.
2. Alm, E. & Baker, D. (1999). Prediction of protein-folding mechanisms from free-energy landscapes derived from native structures. *Proc. Natl Acad. Sci. USA,* **96**, 11305-11310.
3. Debe, D. A. & Goddard, W. A., III (1999). First principles prediction of protein folding rates [In Process Citation]. *J. Mol. Biol.* **294**, 619-625.
4. Galzitskaya, O. V. & Finkelstein, A. V. (1999). A theoretical search for folding/unfolding nuclei in three-dimensional protein structures. *Proc. Natl Acad. Sci. USA,* **96**, 11299-11304.
5. Muñoz, V. & Eaton, W. A. (1999). A simple model for calculating the kinetics of protein folding from three-dimensional structures. *Proc. Natl Acad. Sci. USA,* **96**, 11311-11316.
6. Chiti, F., Taddei, N., White, P. M., Bucciantini, M., Magherini, F., Stefani, M. & Dobson, C. M. (1999). Mutational analysis of acylphosphatase suggests the importance of topology and contact order in protein folding. *Nature Struct. Biol.* **6**, 1005-1009.
7. Martinez, J. C. & Serrano, L. (1999). The folding transition state between SH3 domains is conformationally restricted and evolutionarily conserved. *Nature Struct. Biol.* **6**, 1010-1016.
8. Riddle, D. S., Grantcharova, V. P., Santiago, J. V., Alm, E., Ruczinski, I. & Baker, D. (1999). Experiment and theory highlight role of native state topology in SH3 folding. *Nature Struct. Biol.* **6**, 1016-1024.
9. Grantcharova, V. P., Riddle, D. S., Santiago, J. V. & Baker, D. (1998). Important role of hydrogen bonds in the structurally polarized transition state for folding of the src SH3 domain [see comments]. *Nature Struct. Biol.* **5**, 714-720.

10. Viguera, A. R., Serrano, L. & Wilmanns, M. (1996). Different folding transition states may result in the same native structure. *Nature Struct. Biol.* **3**, 874-880.
11. Martinez, J. C., Pisabarro, M. T. & Serrano, L. (1998). Obligatory steps in protein folding and the conformational diversity of the transition state [see comments]. *Nature Struct. Biol.* **5**, 721-729.
12. Grantcharova, V. P., Riddle, D. S. & Baker, D. (2000). Long-range order in the src SH3 folding transition state [In Process Citation]. *Proc. Natl Acad. Sci. USA,* **97**, 7084-7089.
13. Fersht, A. R. (1995). Characterizing transition states in protein folding: an essential step in the puzzle. *Curr. Opin. Struct. Biol.* **5**, 79-84.
14. Otzen, D. E. & Fersht, A. R. (1998). Folding of circular and permuted chymotrypsin inhibitor 2: retention of the folding nucleus. *Biochemistry,* **37**, 8139-8146.
15. 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.
16. Matouschek, A. & Fersht, A. R. (1993). Application of physical organic chemistry to engineered mutants of proteins: Hammond postulate behavior in the transition state of protein folding. *Proc. Natl Acad. Sci. USA,* **90**, 7814-7818.
17. Clementi, C., Jennings, P. A. & Onuchic, J. N. (2000). How native-state topology affects the folding of dihydrofolate reductase and interleukin-1beta. *Proc. Natl Acad. Sci. USA,* **97**, 5871-71876.
18. Clementi, C., Nymeyer, H. & Onuchic, J. N. (2000). Topological and energetic factors: what determines the structural details of the transition state ensemble and ''Enroute'' intermediates for protein Folding? An investigation for small globular proteins. *J. Mol. Biol.* **298**, 937-953.
19. Onuchic, J. N., Nymeyer, H., Garcia, A. E., Chahine, J. & Socci, N. D. (2000). The energy landscape theory of protein folding: insights into folding mechanisms and scenarios. *Advan. Protein Chem.* **53**, 87-152.
20. McCallister, E. L., Alm, E. & Baker, D. (2000). Critical role of beta-hairpin formation in protein G folding. *Nature Struct. Biol.* **7**, 669-673.
21. Kim, D. E., Fisher, C. & Baker, D. (2000). A breakdown of symmetry in the folding transition state of protein L. *J. Mol. Biol.* **298**, 971-984.
22. Timasheff, S. N. & Arakawa, T. (1989). Stabilization of protein structure by solvents. In *Protein Structure: A Practical Approach* (Creighton, T. E., ed.), pp. 331-335, IRL Press, Oxford.
23. Plaxco, K. W., Simons, K. T. & Baker, D. (1998). Contact order, transition state placement and the refolding rates of single domain proteins. *J. Mol. Biol.* **277**, 985-994.
24. Falzone, C. J., Kao, Y. H., Zhao, J., Bryant, D. A. & Lecomte, J. T. (1994). Three-dimensional solution structure of PsaE from the cyanobacterium Synechococcus sp. strain PCC 7002, a photosystem I protein that shows structural homology with SH3 domains. *Biochemistry,* **33**, 6052-6062.
25. Moran, L. B., Schneider, J. P., Kentsis, A., Reddy, G. A. & Sosnick, T. R. (1999). Transition state heterogeneity in GCN4 coiled coil folding studied by using mutlisite mutations and crosslinking. *Proc. Natl Acad. Sci. USA,* **96**, 10699-10704.
26. Zitzewitz, J. A., Ibarra-Molero, B., Fishel, D. R., Terry, K. L. & Matthews, C. R. (2000). Preformed secondary structure drives the association reaction of GCN4- p1, a model coiled-coil system. *J. Mol. Biol.* **296**, 1105-1116.
27. Grantcharova, V. P. & Baker, D. (1997). Folding dynamics of the src SH3 domain. *Biochemistry,* **36**, 15685-15692.
28. Kraulis, P. J. (1991). MOLSCRIPT: A Program to Produce Both Detailed and Schematic Plots of Protein Structures. *J. Appl. Crystallog.* **24**, 946-950.
29. Bacon, D. J. & Anderson, W. F. (1988). A fast algorithm for rendering space-filling molecule pictures. *J. Mol. Graph.* **6**, 219-220.
30. Merritt, E. A. & Bacon, D. J. (1997). Raster3D: Photorealistic Molecular Graphics. *Methods Enzymol.* **277**, 505-524.

Edited by C. R. Matthews

(Received 27 July 2000; received in revised form 16 November 2000; accepted 21 November 2000)