Long-range order in the src SH3 folding transition state

Viara P. Grantcharova, David S. Riddle*, and David Baker†

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

Communicated by Robert L. Baldwin, Stanford University Medical Center, Stanford, CA, April 19, 2000 (received for review February 21, 2000)

One of the outstanding questions in protein folding concerns the degree of heterogeneity in the folding transition state ensemble: does a protein fold via a large multitude of diverse ''pathways,'' or are the elements of native structure assembled in a well defined order? Herein, we build on previous point mutagenesis studies of the src SH3 by directly investigating the association of structural elements and the loss of backbone conformational entropy during folding. Double-mutant analysis of polar residues in the distal β-hairpin and the diverging turn indicates that the hydrogen bond network between these elements is largely formed in the folding transition state. A 10-glycine insertion in the n-src loop (which connects the distal hairpin and the diverging turn) and a disulfide crosslink at the base of the distal β-hairpin exclusively affect the folding rate, showing that these structural elements are nearly as ordered in the folding transition state as in the native state. In contrast, crosslinking the base of the RT loop or the N and C termini dramatically slows down the unfolding rate, suggesting that dissociation of the termini and opening of the RT loop precede the rate-limiting step in unfolding. Taken together, these results suggest that essentially all conformations in the folding transition state ensemble have the central three-stranded β-sheet formed, indicating that, for the src homology 3 domain, there is a discrete order to structure assembly during folding.

One of the major differences between the "old" and the "new" views of protein folding is the degree of heterogeneity in the folding transition state ensemble (1). In the limit of a single folding pathway, all folding trajectories are presumed to undergo similar conformational transitions, whereas in the limit of a perfectly symmetric folding "funnel," a vast number of different trajectories lead to the native state. Studies of simple lattice models have suggested both "single" and "multiple" folding nuclei scenarios (2, 3). For small proteins that fold without detectable intermediates, characterization of the folding transition state by mutational analysis is perhaps the best available approach to addressing this issue. This method, pioneered by Fersht and coworkers (4, 5), has proven extremely powerful in providing site-specific information about structure at the rate-limiting step for folding (6–11). It probes the formation of side chain–side chain interactions in the transition state by deleting parts of individual residues and assessing the effect on folding kinetics. There are, however, two shortcomings of this approach: (*i*) residues are often involved in multiple interactions, and point mutagenesis does not distinguish which of these are important in the transition state; and (*ii*) the conformation of the peptide backbone can be deduced only indirectly. To go beyond these limitations, in the present study, we employ double-mutant analysis to probe side chain–side chain interactions between structural elements and a 10-glycine insertion and disulfide crosslinks to test backbone ordering and the association of entire structural elements at the transition state.

Previously, we studied the structure of the transition state for folding of the 57-residue src SH3 domain by characterizing the kinetic consequences of a large number of point mutations and found that the rate-limiting step in folding involves formation of the distal β-hairpin and the diverging turn (Fig. 1*A*; refs. 12 and 13). In this study, we investigate long-range order in the transition state by: (*i*) double-mutant analysis of a hydrogen bond network between the distal β-hairpin and the diverging turn to probe their association in the transition state; (*ii*) a 10-glycine insertion in the n-src loop to investigate its conformational rigidity and thus the association of the distal β-hairpin and the diverging turn, which the n-src loop connects; (*iii*) disulfide crosslinking the distal β-hairpin and the RT loop to probe the extent of closure of the two hairpin loops; and (*iv*) disulfide crosslinking of the N and C termini to probe the association of the terminal strands in the transition state.

Methods

Mutagenesis. Point mutagenesis was accomplished with the Quick Change Site-Directed mutagenesis kit (Stratagene). The glycine insertion mutant was constructed by using PCR cassette mutagenesis with primers coding for the 10 glycines. Plasmids harboring the mutations were transformed into BL21 cells, and protein was overexpressed and purified (12). The His tag was not removed for the purposes 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 protein identity.

Disulfide Crosslinking. Residues mutated to cysteine were chosen to satisfy the geometric requirements for disulfide bond formation: Cα–Cα (4.4–6.8 Å), Cβ–Cβ (3.5–4.5 Å), and dihedral angle close to 90° (14). The residues chosen were previously determined to have a very small or no effect on the rate of folding and stability (13). W43C is the only mutation that is likely to affect stability significantly (as judged from the W43A mutation, $\Delta\Delta G$ = 1.2 kcal/mol). Disulfide bonds were oxidized in the presence of 20 mM $K_3Fe(CN)_6$ for 10 min at room temperature. Reactions were performed in the dark because $K_3Fe(CN)_6$ is light sensitive. Disulfide formation was confirmed with Ellman's reagent.

Biophysical Analysis. Protein solutions (100 μM) were made in 50 mM sodium phosphate (pH 6), and the temperature was held constant at 295 K. For wild type (WT) and the S47A mutant, experiments were also performed in 50 mM NaPi (pH 3) at 295 K. To reduce the disulfide-crosslink mutants, they were incubated in 10 mM DTT for 1 h, and the same concentration of reducing agent was present throughout the kinetic experiments. The kinetics of folding and unfolding were followed by tryptophan fluorescence on a Bio-Logic SFM-4 stopped-flow instrument (Molecular Kinetics, Pullman, WA). The unfolding reaction for the WT protein can be modeled as a two-state process (15), and the kinetic and equilibrium data for the mutants were

Abbreviations: SH3, src homology 3; WT, wild type; Gnd, guanidine.

*Present address: Department of Immunology, Mayo Clinic, Rochester, MN 55904

†To whom reprint requests should be addressed. E-mail: dabaker@u.washington.edu.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

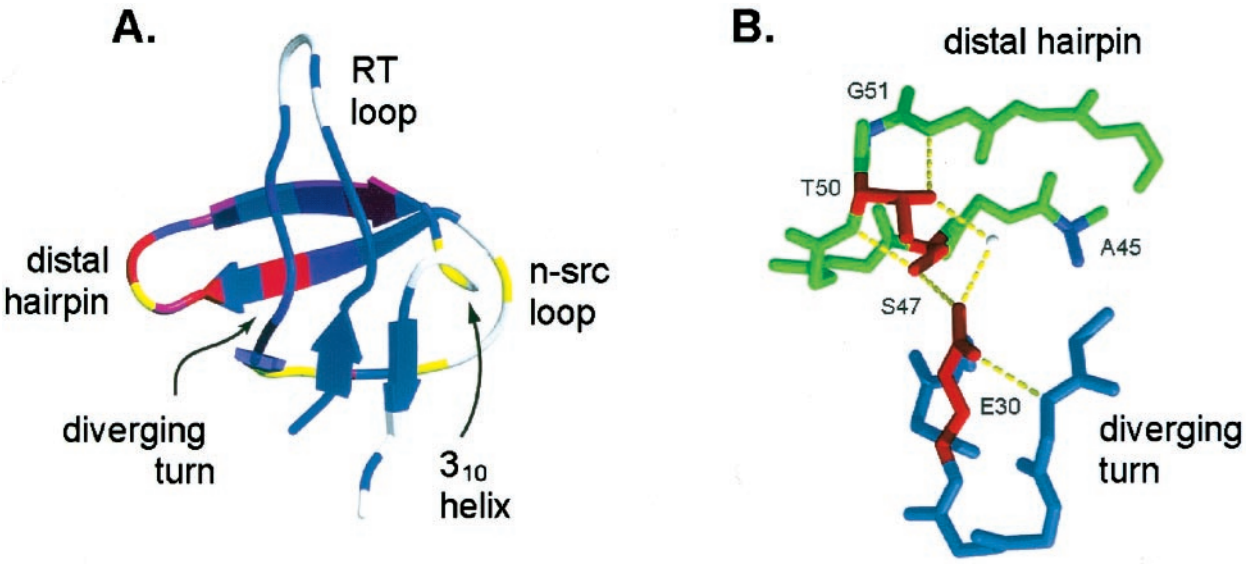

Fig. 1. (*A*) Structure of the src SH3 domain (1fmk.pdb) colored by previously reported Φ_F-values (13) on a continuous scale from red ($\Phi_F = 1$) to blue ($\Phi_F = 0$). Residues at which mutations increase or decrease both k_f and k_u are colored in yellow. The graphic was generated with MOLSCRIPT (33) and RASTER3D (34, 35). (*B*) Structure of the hydrogen bond network between the β-distal hairpin and the diverging turn (MIDAS; refs. 36 and 37). Residues included in the double-mutant cycles are shown in red.

fit to a two-state model. For each mutant, the free energy of folding is calculated as:

$$\Delta G_{U\text{-}F} = RT\ln(k_f/k_u), \quad [1]$$

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

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

$$\Delta\Delta G_{U\text{-}\ddagger} = RT\ln(k_f^{wt}/k_f^{mut}), \quad [2]$$

where k_f and k_u are the rates of folding and unfolding, respectively, at denaturant concentrations experimentally accessible for that mutant. The parameter Φ_F is defined as

$$\Phi_F = \Delta\Delta G_{U\text{-}\ddagger}/\Delta\Delta G_{U\text{-}F} \quad [3]$$

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 (5). In the case of the double mutants, a Φ_F-value for the pairwise interaction can be determined similarly:

$$\Phi_F^{int} = \Delta\Delta G_{U\text{-}\ddagger}^{int}/\Delta\Delta G_{U\text{-}F}^{int}. \quad [4]$$

Loop Entropy Estimates. The change in the free energy of the unfolded state as a result of loop insertion or disulfide crosslinking can be estimated from polymer theory (16):

$$\Delta G = -RT\ln(L/L_o), \quad [5]$$

where L_o and L are loop lengths before and after the modification, respectively. In the case of the 10-glycine insertion in the n-src loop, the original loop length is 5 and after the insertion it is 15. In the case of the three disulfide crosslinks, which generate a loop in the protein that did not exist previously, L_o is taken to be 1, and L is the length of the loop enclosed by the crosslink.

Results

Hydrogen Bond Network Between the Distal β-Hairpin and the Diverging Turn. Mutagenic analysis of the SH3 domain folding transition state revealed the clustering of structured residues in the distal β-hairpin and the diverging turn (12). Mutagenesis also suggested that these elements might interact with each other at the rate-limiting step, because mutations in the hydrogen bond network between the distal β-hairpin and the diverging turn (Fig. 1*B*) had a dramatic effect on the rate of folding. In this study, we performed double-mutant cycles on these hydrogen bond network residues (S47 and T50 in the distal β-hairpin and E30 in the diverging turn) to quantify interaction energies at the transition state (4). Both E30A_S47A (Fig. 2*A*) and E30A_T50A (Fig. 2*B*) double mutants are considerably less destabilized than expected from the sum of the single-mutant effects (Table 1): in the native state, the interaction energy between the two mutated residues is 1.02 kcal/mol for E30A and S47A (Fig. 2*D*) and 1.08 kcal/mol for E30A and T50A. A large fraction of this interaction energy is present at the folding transition state: 0.78 kcal/mol for E30 and S47 and 0.83 kcal/mol for E30 and T50, yielding interaction Φ_F-values of 0.76 and 0.77, respectively. The finding that these hydrogen bonds are already formed at the transition state confirms directly the association of the distal β-hairpin and the diverging turn in the transition state inferred from the point mutagenesis experiments.

The E30A mutation removes a large portion of this buried residue and is therefore expected to be quite disruptive. As a less drastic probe of the interactions in the hydrogen bond network, we compare the kinetic effect of the S47A mutation at pH 3 and pH 6 (Fig. 2*C* and Table 1). At pH 3, the carboxyl group of E30 probably is partially protonated (depending on the local pKa), thus disrupting some of its interactions with the distal β-hairpin (S47). The effect of S47A on stability and the rate of folding is smaller at pH 3 than at pH 6, but the Φ_F-value is still 1. This effect is consistent with the idea that some of the interactions between S47 and E30 present at the transition state at pH 6 can be disrupted by low pH.

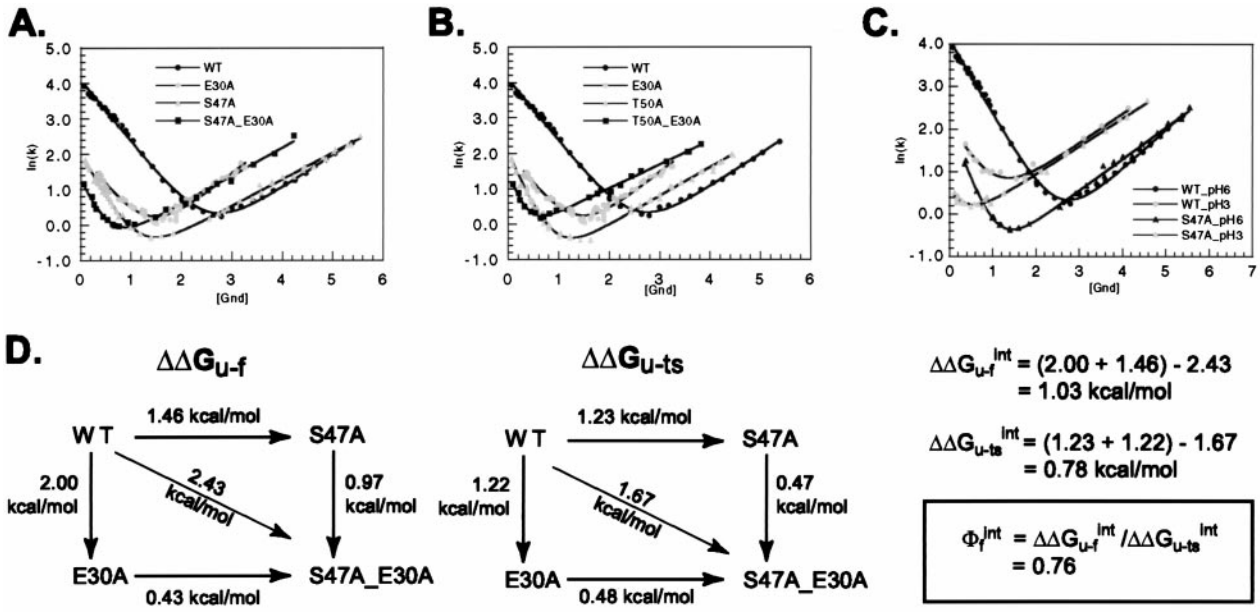

Fig. 2. Kinetic analysis of E30A_S47A (*A*) and E30A_T50A (*B*) in 50 mM NaPi, pH 6, at 295 K. Gnd, guanidine. (*C*) Kinetic analysis of WT and S47A mutant in 50 mM NaPi, pH 3, at 295 K. Solid lines indicate the best fit to the data with KALEIDAGRAPH. (*D*) Double-mutant cycle analysis to determine the interaction energy between E30A and S47A in the native ($\Delta\Delta G_{U\text{-}F}$) and transition states ($\Delta\Delta G_{U\text{-}\ddagger}$). Data for the calculations are presented in Table 1. $\Delta\Delta G_{U\text{-}\ddagger}$ is calculated as $RT\ln(k_f^{wt}/k_f^{mut})$. Energies for the E30A_T50A double mutant are not shown explicitly but can be calculated similarly by using data from Table 1.

Backbone Restriction in the Folding Transition State. In addition to manipulating interaction energies of side chains by point mutagenesis, we can alter the entropy of the protein backbone by engineering loop insertions and covalent crosslinks and then use the effect of such changes on the folding kinetics as a reporter on the degree of association of the structural elements in the transition state. As in Φ-value analysis, the premise of these experiments is that changes in the rates of folding and unfolding will reveal the extent to which two elements are brought together in the transition state compared with the denatured and native states. To facilitate interpretation of the kinetic results, it is reasonable to assume that the primary effect of these modifications is on chain conformational entropy. The denatured state, which is most disordered, is expected to be affected significantly by the modifications: glycine insertion in loops increases the entropy of the denatured state and thus lowers its free energy, whereas disulfide crosslinks decrease the entropy of the denatured state and destabilize it in proportion to the length of the crosslinked fragment (Fig. 3). The entropy of the native state, on the other hand, should not change greatly in the case of the disulfide crosslinks, because the elements we are probing already interact fully in the native state. In the case of the 10-glycine mutant, the entropy of the native state will increase but to a lesser extent than the entropy of the denatured state. There might be some destabilization of the native state resulting from disruption of local interactions at the site of glycine insertion or strain from suboptimal disulfide geometry; however, we have tried to minimize these effects by choosing the modification sites carefully. The entropic effect on the transition state then depends on the proximity of the structural elements being probed and can be deduced from the changes in k_f and k_u. If only k_f is affected, it can be concluded that the crosslinked regions are as ordered in the transition state as in the native state (Fig. 3 *A* and *B*), whereas, if only k_u changes, the region is likely to be as disordered in the transition state as in the denatured state.

Table 1. Kinetic parameters for the hydrogen bond network mutants

Mutant	$\ln(k_f)^{0M}$	$\ln(k_u)^{3M}$	m_f	m_u	ΔG_{U-F}	$\Delta\Delta G_{U-F}$
WT*	4.10	0.139	1.02	0.54	3.95	—
E30A*	2.03	1.48	1.09	0.65	2.28	−2.00
S47A*	2.01	0.537	1.50	0.44	2.18	−1.46
T50A*	1.99	0.750	1.84	0.47	2.14	−1.60
E30A_S47A	1.21	1.39	†	0.46	1.28	−2.43
E30A_T50A	1.32	1.66	†	0.39	0.977	−2.52
WT_pH3	2.08	1.69	1.08	0.41	1.46	−2.09
S47A_pH3	0.052	1.59	1.45	0.38	0.249	−3.23

Kinetics of folding and unfolding were followed by changes in tryptophan fluorescence on a stopped flow instrument at 295 K; k_f is reported in the absence of denaturant, and k_u is in 3 M Gnd to avoid extrapolation; m_f and m_u are the dependences of the folding and the unfolding rates, respectively, on Gnd. ΔG_{U-F} (free energy of unfolding) and $\Delta\Delta G_{U-F}$ (the difference in ΔG_{U-F} between WT and the mutant proteins) were calculated from the kinetic parameters as described in the *Methods* section. Typical errors for the kinetic measurements are 1–10% as reported in ref. 13.

*Kinetic data for these mutants were published previously in ref. 12.

†These values could not be estimated reliably because of the small region over which k_f can be measured.

Glycine Insertion in the n-src Loop. The involvement of the n-src loop in the transition state was difficult to evaluate from the point mutagenesis analysis because of the very small effect on stability of all of the mutations in this region (13). An insertion of 10 glycine residues in the n-src loop between residues 40 and 41 (Fig. 3*A*) was designed to test the importance of this loop in bringing the distal β-hairpin and the diverging turn together at the transition state. The addition of 10 glycine residues considerably increases the entropic cost of bringing these two elements together and is expected to decrease the rate of folding if association of the two elements is required in the transition state. A correlation between the rate of folding and the length of this loop was observed in a comparison of homologous SH3 domains (17): the phosphatidylinositol 3-kinase SH3 domain has the longest n-src loop, and its folding rate is the slowest of the SH3 domains that have been characterized (18). Furthermore, in a

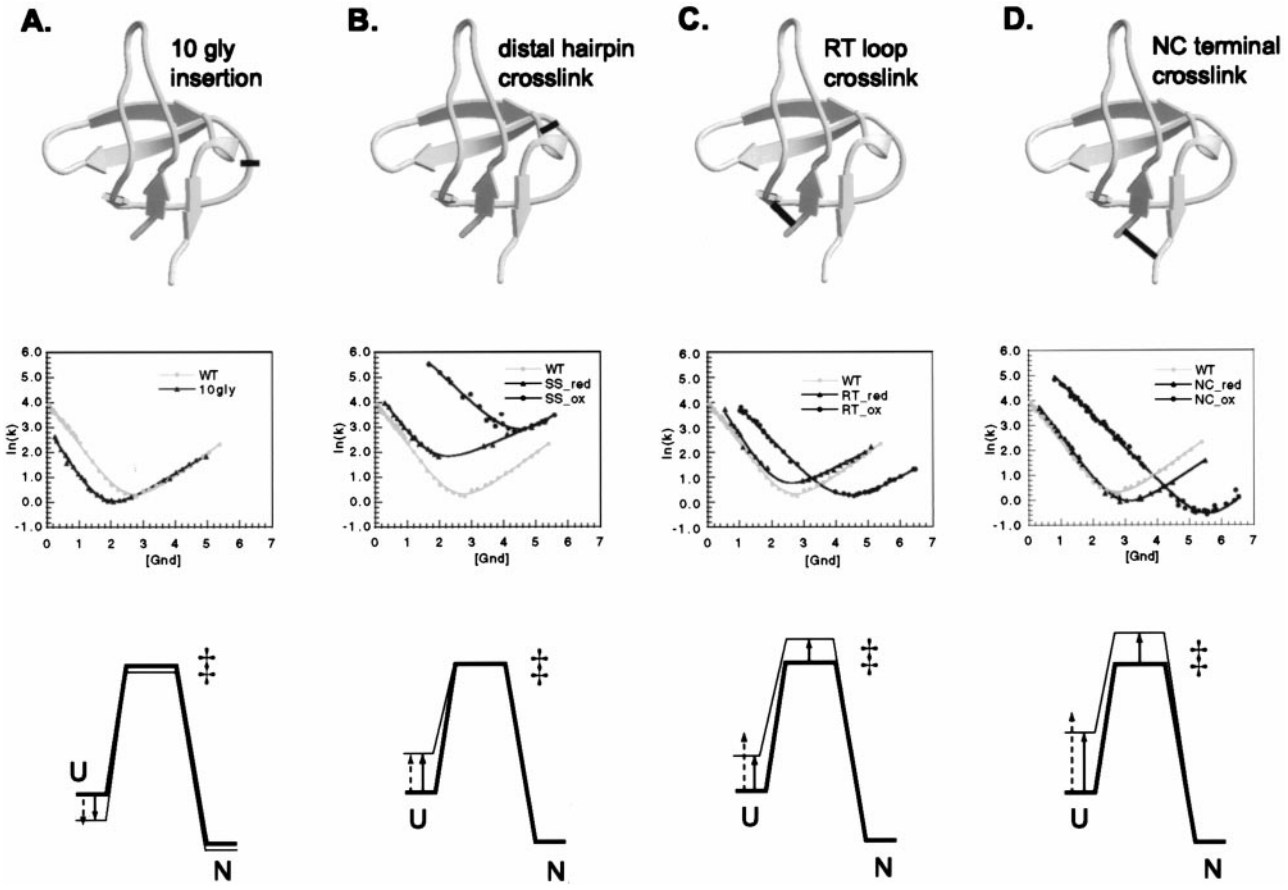

Fig. 3. Characterization of the 10-glycine insertion (*A*), the distal hairpin crosslink (*B*), the RT loop crosslink (*C*), and the N- and C-terminal crosslink mutants (*D*). For each modification, panels from *Top* to *Bottom* show site of the modification, kinetic analysis, and an energy diagram explaining the effects on kinetics. The structure diagrams of the SH3 domain were generated with MOLSCRIPT (33) and RASTER3D (34, 35). Energy diagrams show the relative free energies of the unfolded (U), the transition (‡), and the native state (N) before (thick lines) and after (thin lines) each modification. In the case of the 10-glycine insertion, the reference protein is the WT, and, for the three disulfide crosslinks, the reference is the reduced version of the particular double cysteine mutant. Solid arrows indicate the experimentally determined changes in the free energy of denatured state, and dashed arrows indicate the changes predicted from loop entropy estimates (16). 10gly, 10-glycine; NC, N- and C-terminal crosslink; SS, distal hairpin crosslink; red, reduced; ox, oxidized.

combinatorial mutagenesis experiment on the SH3 domain, phage-display selection of correctly folded proteins yielded an enrichment for shortened n-src loops (19). The 10-glycine insertion leads to an overall decrease in stability of 0.72 kcal/mol (Table 2), consistent with loop entropy estimates ($-1.5RT\ln(L/L_o) = 0.78$ kcal/mol, where L_o and L are the original and the new loop lengths, respectively; ref. 16). Remarkably, the kinetic effect of the insertion (Fig. 3*A* and Table 2) is exclusively on the rate of folding, k_f. The structural elements on either end of the n-src loop (the distal β-hairpin and the diverging turn) seem to be fully associated at the folding transition state, consistent with the double-mutant results (Fig. 2). The lack of effect of such a large insertion on the unfolding rate, k_u, is striking and suggests that unfolding is initiated by the dissociation of the two terminal strands and the RT loop but does not involve disruption of the n-src loop.

Crosslinking the Distal β-Hairpin. Φ-Value analysis of both the src and the α-spectrin SH3 domains highlighted the importance of the distal β-hairpin in the folding transition state (12, 20). For both proteins, the tip of the distal β-hairpin is the only region that contains residues with Φ_F-values equal to 1 (in the SH3 domain, S47, T50, and G51 have all of their local interactions formed at the transition state). The middle of the hairpin, however, appeared somewhat flexible (L44A and Y55A on the solvent-exposed side of the hairpin have intermediate Φ-values), and there were no suitable mutations to probe strand association at the base of the hairpin. To examine specifically the backbone

Table 2. Kinetic parameters for 10-glycine insertion and disulfide crosslink mutants

Mutant	$\ln(k_f)^{1M}$	$\ln(k_u)^{5.5M}$	m_f	m_u	ΔG_{U-F}	$\Delta\Delta G_{U-F}$
WT	2.36	2.44	1.02	0.54	3.95	—
10gly	0.997	2.30	1.20	0.44	2.86	−0.72
SS_red	2.75	3.38	1.05	0.33	2.50	−0.32
SS_ox	6.50	3.37	0.750	0.59	5.83	1.88
RT_red	2.65	2.37	1.18	0.39	3.47	0.21
RT_ox	3.81	0.62	0.78	0.42	4.98	1.92
NC_red	2.58	1.59	1.03	0.50	4.36	0.63
NC_ox	4.73	−1.30	0.766	0.80	8.70	3.58

Kinetics of folding and unfolding were followed by changes in tryptophan fluorescence on a stopped flow instrument at 295 K; k_f is reported in 1 M Gnd, and k_u is in 5.5 M Gnd to avoid extrapolation; m_f and m_u are the dependences of the folding and the unfolding rates, respectively, on Gnd. ΔG_{U-F} (free energy of unfolding) and $\Delta\Delta G_{U-F}$ (the difference in ΔG_{U-F} between the particular mutant protein and WT) were calculated from the kinetic parameters as described in the *Methods* section. Typical errors for the kinetic measurements are 1–10% as reported in ref. 13. Abbreviations are defined in the legend to Fig. 3.

conformation of the β-hairpin in the transition state, we tested the kinetic consequences of covalently crosslinking the hairpin. For this purpose, W43 and S58 at the base of the distal β-hairpin (Fig. 3*B*) were mutated to cysteines and then crosslinked by using an oxidizing agent (see *Methods* section). Under reducing conditions, the double-cysteine mutant (SS mutant) is destabilized compared with the WT SH3 domain; however, oxidation significantly stabilizes the mutant protein: $\Delta\Delta G_{U\text{-}F}$ between the reduced and oxidized forms of the mutant is 2.55 kcal/mol. This value matches closely the theoretical loop entropy estimate from polymer theory (2.38 kcal/mol; see *Methods* section; ref. 16), suggesting that stabilization results largely from the decrease in entropy of the denatured state. Kinetic analysis reveals that the oxidized protein folds 30 times faster than the reduced form, whereas the unfolding rate is virtually unchanged (Fig. 3*B* and Table 2). The resulting Φ_F-value of 1 for the disulfide crosslink unambiguously confirms that the distal β-hairpin is conformationally restricted at the transition state (Fig. 3*B*). As in the case of the glycine insertion, the lack of effect of the crosslink on k_u is remarkable: the unfolding event must not involve even partial unraveling of the distal β-hairpin. Flexibility in the middle of the hairpin therefore does not prevent the conformational locking of the hairpin's base.

Crosslinking the RT Loop. To probe the extent of formation of the RT loop in the folding transition state, we have introduced a disulfide crosslink at its base. T9 at the N terminus of the protein and Q33 after the diverging turn were mutated to cysteine and oxidized to close off a loop of 23 residues covalently (Fig. 3*C*). Oxidation stabilizes the protein by 1.71 kcal/mol, suggesting that the disulfide bond geometry is favorable and does not introduce strain in the native state of the protein. This stabilization, however, is less than the loop entropy reduction estimate (2.75 kcal/mol; see *Methods* section; ref. 16), suggesting that the RT loop might be partially structured in the denatured state. Kinetic analysis (Fig. 3*C* and Table 2) shows that, in contrast to the distal hairpin crosslink, formation of the RT loop crosslink dramatically decreases the unfolding rate, suggesting that the rate-limiting step in unfolding involves the opening of the RT loop. Crosslinking also increases the folding rate, indicating that parts of the RT loop might be structured in the transition state. Considering the point mutagenesis results that revealed Φ-values uniformly close to 0 throughout the N-terminal strand and the tip of the RT loop, it is most likely that the crosslink itself has caused an expansion of the structured region of the transition state to include the RT loop. Another possibility is that the RT loop is stabilized primarily by backbone hydrogen bonds and not by side chain–side chain interactions in the transition state.

Crosslinking the N and C Termini. The point mutagenesis analysis suggested that the N and C termini do not associate at the folding transition state, because none of the mutations in this region affect the rate of folding. We probed the association between the N- and C-terminal strands in the transition state by engineering a disulfide crosslink between them. Two cysteine mutations were introduced in the SH3 domain (T9C at the N terminus and S64C at the C terminus) and then crosslinked as described for the distal hairpin. Comparison of the T9C_S64C mutant (NC mutant) under reducing and oxidizing conditions (Fig. 3*D* and Table 2) reveals that the oxidized protein is stabilized significantly ($\Delta\Delta G_{U\text{-}F}$ = 2.8 kcal/mol) but less than expected from the effect of crosslinking on the entropy of the denatured state ($\Delta\Delta G$ = 3.52 kcal/mol; see *Methods* section; ref. 16; the denatured state may be more ordered than the random coil model assumed in the loop entropy estimate). Kinetic measurements show that both the folding and the unfolding rates of the SH3 domain are affected roughly equally by the disulfide crosslink. In general, circularization is always expected to increase k_f because of the greater decrease in the entropy of the denatured state compared with that of the transition state, but k_u will decrease only if the termini are apart at the transition state. (In proteins, like acyl-CoA-binding protein, whose termini interact at the transition state, crosslinking would be expected to affect primarily the rate of folding; ref. 7.) The decrease in the unfolding rate brought about by the NC crosslink suggests that the termini are not as ordered in the SH3 transition state as they are in the native state.

Discussion

SH3 Folding Transition State. Φ-Value analysis has become the method of choice for studying folding transition states (4, 5). In this study, we have extended the repertoire of probes of the transition state ensemble to include glycine insertion and disulfide crosslinking as direct experimental measures of the conformational constraints on the peptide backbone and the association of structural elements. Engineering of disulfide crosslinks has been used to assess the effect of reducing conformational entropy in different parts of the molecule (21–25) or to explore transition-state heterogeneity (26, 27). Loop insertions have been used to explore the role of loops in determining protein stability and the folding mechanism (28–30).

Our current findings from the double-mutant analysis, 10-glycine insertion, and covalent crosslinking combined with the point mutagenesis results provide a comprehensive picture of the folding transition state for the SH3 domain. The distal β-hairpin still stands out as the structural element best formed in the transition state; however, now we have information about strand pairing along the entire length of the hairpin. Previous experiments had indicated that the tip of the hairpin is well ordered as judged by the clustering of high Φ-value residues; however, the middle of the hairpin is likely not as rigid in the transition state as in the native state, because solvent-exposed residues are paired only partially with each other. The finding that the distal β-hairpin crosslink increases the rate of folding and has no effect on the unfolding rate suggests that the two strands come in close proximity at the base of the hairpin before the folding transition state. Interestingly, similar results were found for one of the hairpins in protein L (38), suggesting that "looping" in the middle of the hairpin might be a common theme during protein folding. Constraining the tip and the base of the hairpin may be important for specifying the topology of the folding protein, whereas keeping the middle flexible would allow hydrophobic core rearrangements after the transition state.

The current experiments establish firmly the interaction of the distal β-hairpin and the diverging turn in the transition state. The double-mutant results provide concrete evidence that the hydrogen bond network between these two elements is mostly formed at the transition state. The high Φ_F-values for these nonlocal hydrogen bond interactions (0.78 and 0.83) indicate that most of the interaction energy is present at the rate-limiting step; further alignment of the hydrogen bond geometries and immobilization of the participating residues after the transition state must contribute the remaining ≈20%. It is generally assumed that transition states for folding are stabilized primarily by hydrophobic interactions, with hydrogen bonds contributing only later to the stability of the native state because of their stricter geometric requirements. The SH3 domain is the first case in which nonlocal side chain hydrogen bonds have been found to stabilize the transition state (31). The hydrogen bond network does not, however, seem to be required for folding (e.g., to confer specificity or determine the alignment of structural elements). In the α-spectrin SH3 domain, for example, this interaction is replaced by a hydrophobic cluster, and in the fyn SH3 domain (78% homologous to the SH3 domain), the same hydrogen bond network is in place in the native state but does not contribute to stabilization of the transition state (A. Davidson, personal communication). The difference between the src and fyn SH3

domains may be a result of the slightly different structures of their diverging turns: perhaps the two large phenylalanines in fyn (instead of Phe and Leu in src) cannot pack closely next to each other and interact with the distal β-hairpin until after the transition state. The local variation in the types of interactions stabilizing the transition state is an illustration of the lack of selective pressure on the details of the protein-folding mechanism (32).

The association of the distal β-hairpin and the diverging turn at the transition state is confirmed further by the 10-glycine insertion into the n-src loop connecting the two elements. The exclusive effect of the insertion on the folding rate strongly suggests that the peptide backbone of the regions flanking the loop is constricted at the rate-limiting step. In marked contrast, all previous loop-insertion experiments revealed that both folding rates decrease and unfolding rates increase as loops lengthen. The context dependence of loop lengthening has been noted before; however, it has been explained mostly in terms of the flexibility of the region in the native state (28–30). Our findings indicate that the kinetic effect of loop elongation is related directly to the extent to which the elements connected by the loop are topologically constrained at the transition state.

The effects of the disulfide crosslinks and the loop insertion on the unfolding rate highlight the pronounced hierarchy of events. The distal β-hairpin crosslink and the glycine loop insertion have no effect on k_u, clearly indicating that the three-stranded sheet formed by the interaction of the distal β-hairpin and the diverging turn remains intact at the unfolding transition state. In contrast, the RT loop crosslink and the NC crosslink dramatically slow down the unfolding rate, suggesting that the rate-limiting step in unfolding involves the dissociation of the N and C termini and the opening of the RT loop. These results indicate an even greater structural polarization of the folding transition state than suggested by our previous studies.

The conformational restriction of structural elements in the src SH3 domain transition state has implications for the mechanism of folding. The transition state ensemble consists of a relatively small number of conformers, all of which have the distal β-hairpin and the diverging turn ordered and interacting with each other, and with heterogeneities limited to the RT loop and the N and C termini. The energy landscape of this protein, therefore, significantly deviates from a symmetrical funnel in which the transition state includes all conformations with a particular degree of freedom. The folding of the src SH3 domain is surprisingly consistent with the more traditional single pathway-based picture of protein folding.

We thank Andreas Matouschek for providing us with his disulfide-crosslinking protocol. We are grateful to members of the Baker group for their useful comments on the manuscript. This work was supported by grants from the National Institutes of Health and the Office of Naval Research and a Young Investigator award to D.B. from the Packard Foundation.

1. Pande, V. S., Grosberg, A., Tanaka, T. & Rokhsar, D. S. (1998) *Curr. Opin. Struct. Biol.* **8,** 68–79.
2. Shakhnovich, E. I. (1998) *Folding Des.* **3,** R108–R111.
3. Abkevich, V. I., Gutin, A. M. & Shakhnovich, E. I. (1994) *Biochemistry* **33,** 10026–10036.
4. Fersht, A. R., Matouschek, A. & Serrano, L. (1992) *J. Mol. Biol.* **224,** 771–782.
5. Fersht, A. R. (1995) *Curr. Opin. Struct. Biol.* **5,** 79–84.
6. Chiti, F., Taddei, N., White, P. M., Bucciantini, M., Magherini, F., Stefani, M. & Dobson, C. M. (1999) *Nat. Struct. Biol.* **6,** 1005–1009.
7. Kragelund, B. B., Osmark, P., Neergaard, T. B., Schiodt, J., Kristiansen, K., Knudsen, J. & Poulsen, F. M. (1999) *Nat. Struct. Biol.* **6,** 594–601.
8. López-Hernández, E. & Serrano, L. (1995) *Folding Des.* **1,** 43–55.
9. Main, E. R., Fulton, K. F. & Jackson, S. E. (1999) *J. Mol. Biol.* **291,** 429–444.
10. Martinez, J. C. & Serrano, L. (1999) *Nat. Struct. Biol.* **6,** 1010–1016.
11. Milla, M. E., Brown, B. M., Waldburger, C. D. & Sauer, R. T. (1995) *Biochemistry* **34,** 13914–13919.
12. Grantcharova, V. P., Riddle, D. S., Santiago, J. V. & Baker, D. (1998) *Nat. Struct. Biol.* **5,** 714–720.
13. Riddle, D. S., Grantcharova, V. P., Santiago, J. V., Alm, E., Ruczinski, I. & Baker, D. (1999) *Nat. Struct. Biol.* **6,** 1016–1024.
14. Thornton, J. M. (1981) *J. Mol. Biol.* **151,** 261–287.
15. Grantcharova, V. P. & Baker, D. (1997) *Biochemistry* **36,** 15685–15692.
16. Jacobsen, H. & Stockmayer, W. H. (1950) *J. Chem. Phys.* **18,** 1600–1606.
17. Plaxco, K. W., Guijarro, J. I., Morton, C. J., Pitkeathly, M., Campbell, I. D. & Dobson, C. M. (1998) *Biochemistry* **37,** 2529–2537.
18. Guijarro, J. I., Morton, C. J., Plaxco, K. W., Campbell, I. D. & Dobson, C. M. (1998) *J. Mol. Biol.* **276,** 657–667.
19. Riddle, D. S., Santiago, J. V., Bray-Hall, S. T., Doshi, N., Grantcharova, V. P., Yi, Q. & Baker, D. (1997) *Nat. Struct. Biol.* **4,** 805–809.
20. Martinez, J. C., Pisabarro, M. T. & Serrano, L. (1998) *Nat. Struct. Biol.* **5,** 721–729.
21. Creighton, T. E. (1992) in *Protein Folding*, ed. Creighton, T. E. (Freeman, New York), pp. 301–354.
22. Strausberg, S., Alexander, P., Wang, L., Gallagher, T., Gilliland, G. & Bryan, P. (1993) *Biochemistry* **32,** 10371–10377.
23. Ikeguchi, M., Fujino, M., Kato, M., Kuwajima, K. & Sugai, S. (1998) *Protein Sci.* **7,** 1564–1574.
24. Kobayashi, N., Honda, S. & Munekata, E. (1999) *Biochemistry* **38,** 3228–3234.
25. Zhang, T., Bertelsen, E. & Alber, T. (1994) *Nat. Struct. Biol.* **1,** 434–438.
26. Moran, L. B., Schneider, J. P., Kentsis, A., Reddy, G. A. & Sosnick, T. R. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 10699–10704.
27. Otzen, D. E. & Fersht, A. R. (1998) *Biochemistry* **37,** 8139–8146.
28. Viguera, A. R. & Serrano, L. (1997) *Nat. Struct. Biol.* **4,** 939–946.
29. Ladurner, A. G. & Fersht, A. R. (1997) *J. Mol. Biol.* **273,** 330–337.
30. Nagi, A. D., Anderson, K. S. & Regan, L. (1999) *J. Mol. Biol.* **286,** 257–265.
31. Myers, J. K. & Oas, T. G. (1999) *Biochemistry* **38,** 6761–6768.
32. Plaxco, K. W., Simons, K. T. & Baker, D. (1998) *J. Mol. Biol.* **277,** 985–994.
33. Kraulis, P. J. (1991) *J. Appl. Crystallogr.* **24,** 946–950.
34. Bacon, D. J. & Anderson, W. F. (1988) *J. Mol. Graphics* **6,** 219–220.
35. Merritt, E. A. & Bacon, D. J. (1997) *Methods Enzymol.* **277,** 505–524.
36. Ferrin, T. E., Huang, C. C., Jarvis, L. E. & Langridge, R. (1988) *J. Mol. Graphics* **6,** 13–27.
37. Huang, C. C., Pettersen, E. F., Klein, T. E., Ferrin, T. E. & Langridge, R. (1991) *J. Mol. Graphics* **9,** 230–236.
38. Kim, D. E., Fisher, C. & Baker, D. (2000) *J. Mol. Biol.* **298,** 971–984.