Disulfide Bond-Coupled Folding of Bovine Pancreatic Trypsin Inhibitor Derivatives Missing One or Two Disulfide Bonds†

Phyllis Anne Kosen,*,‡ Cara B. Marks,§,∥ Arnold M. Falick,‡ Stephen Anderson,§,⊥ and Irwin D. Kuntz‡

Department of Pharmaceutical Chemistry, University of California at San Francisco, San Francisco, California 94143, and Department of Cardiovascular Research, Genentech, South San Francisco, California 94080

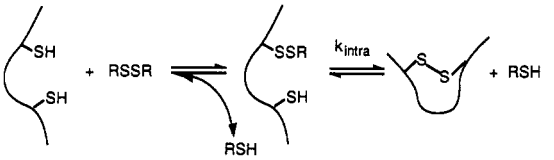
Received September 30, 1991; Revised Manuscript Received March 25, 1992

ABSTRACT: The disulfide bond-coupled folding and unfolding mechanism (at pH 8.7, 25 °C in the presence of oxidized and reduced dithiothreitol) was determined for a bovine pancreatic trypsin inhibitor mutant in which cysteines 30 and 51 were replaced with alanines so that only two disulfides, between cysteines 14 and 38 and cysteines 5 and 55, remain. Similar studies were made on a chemically-modified derivative of the mutant retaining only the 5–55 disulfide. The preferred unfolding mechanism for the Ala30/Ala51 mutant begins with reduction of the 14–38 disulfide. An intramolecular rearrangement via thiol–disulfide exchange, involving the 5–55 disulfide and cysteines 14 and/or 38, then occurs. At least five of six possible one-disulfide bond species accumulate during unfolding. Finally, the disulfide of one or more of the one-disulfide bond intermediates (excluding that with the 5–55 disulfide) is reduced giving unfolded protein. The folding mechanism seems to be the reverse of the unfolding mechanism; the observed folding and unfolding reactions are consistent with a single kinetic scheme. The rate constant for the rate-limiting intramolecular folding step—rearrangements of other one-disulfide bond species to the 5–55 disulfide intermediate—seems to depend primarily on the number of amino acids separating cysteines 5 and 55 in the unfolded chain. The energetics and kinetics of the mutant's folding mechanism are compared to those of wild-type protein [Creighton, T. E., & Goldenberg, D. P. (1984) *J. Mol. Biol. 179*, 497] and a mutant missing the 14–38 disulfide [Goldenberg, D. P. (1988) *Biochemistry 27*, 2481]. The most striking effects are destabilization of the native structure and a large increase in the rate of unfolding.

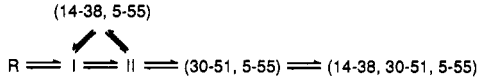
During the course of protein folding and unfolding reactions, disulfide bond-coupled formation or reduction, when mediated by disulfide and thiol reagents, occurs in two steps as diagramed in Scheme I (Creighton, 1975a, 1986; Creighton & Goldenberg, 1984; Snyder, 1987).

The rates of both steps depend on thiol–disulfide exchange chemistry, tempered by solution pH and acidity of the reacting thiol and conjugate thiols of the disulfide (Creighton, 1975a; Szajewski & Whitesides, 1980; Shaked et al., 1980; Snyder et al., 1981; Snyder, 1987).[1] While both steps also depend on steric accessibility and relative orientations of thiol and disulfide, the second step differs from the first in that, in addition, it reflects the collision rate of two polypeptide cysteine-containing regions. It is this second reaction that is equivalent to a conformational transition in a folding reaction not involving disulfide formation. An apparent first-order rate constant for the intramolecular reaction, k_{intra}, can be extracted from observed reaction rate constants (Creighton & Goldenberg, 1984). Folding and unfolding studies, incorporating disulfide formation and reduction, have an advantage that kinetic intermediates can be trapped by alkylation or protonation of any cysteines not involved in a disulfide. Thus, folding steps preceding the slowest observable step are readily characterized and structures of trapped intermediates can be analyzed.

Scheme I



Scheme II



Bovine pancreatic trypsin inhibitor (BPTI;[2] Figure 1), a protein of 58 amino acid residues containing three disulfides, is the paradigm for protein folding reactions coupled to di-

† Supported by Genentech, Inc., by NSF Grant DMB8606901 and NIH Grant GM19267 (to I.D.K.), and by NSF Grant DMB9018707 and NIH Grant R01AG10462 (to S.A.).

* Address correspondence to this author.

‡ University of California at San Francisco.

§ Genentech.

∥ Present address: Medical Research Council Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, England.

⊥ Present address: Center for Advanced Biotechnology and Medicine, Rutgers University, 679 Hoes Lane, Piscataway, NJ 08854.

[1] The reactive species is the thiolate ion. Both steps in Scheme I involve thiolate anions, not thiols. Often, however, these reactions are written as if they were thiol–disulfide exchange reactions (Creighton & Goldenberg, 1984; Goldenberg, 1988). We follow that custom.

[2] Abbreviations: BPTI, bovine pancreatic trypsin inhibitor; Ala30/Ala51 and Ala14/Ala38, mutants of BPTI in which the respective cysteines are replaced by alanines; Ser14/Ser38, a mutant of BPTI in which the respective cysteines are replaced by serines; (Cam14/Cam38; Ala30/Ala51), a chemically-modified derivative of Ala30/Ala51 in which the disulfide formed by cysteines 14 and 38 is reduced selectively and the thiols are blocked by carbamoylmethylation; N, native or fully refolded Ala30/Ala51; R, fully reduced Ala30/Ala51, (Cam14/Cam38; Ala30/Ala51), or BPTI; I, one-disulfide bond intermediates of Ala30/Ala51 excluding the 5–55 disulfide bond intermediates or all of the BPTI one-disulfide bond intermediates (the content clarifies which definition of R or I applies); tetra(carbamoylmethyl)-Ala30/Ala51, fully reduced and carbamoylmethylated Ala30/Ala51, Ala30/Ala51, (Cam14/Cam38; Ala30/Ala51), BPTI, and kinetic intermediates are also referred to by the residue numbers involved in a disulfide. For example, (Cam14/Cam38; Ala30/Ala51) is also called (5–55)); DTT^{S}_{S}, DTT^{SH}_{SH}, dithiothreitol (Cleland's reagent); DTT^{S}_{S}, *trans*-4,5-dihydroxy-1,2-dithiane (oxidized dithiothreitol); Gdn-HCl, guanidine hydrochloride; GSH, reduced glutathione; GSSG, oxidized glutathione; TFA, trifluoroacetic acid; Tris, tris(hydroxymethyl)aminomethane.

0006-2960/92/0431-5705$03.00/0

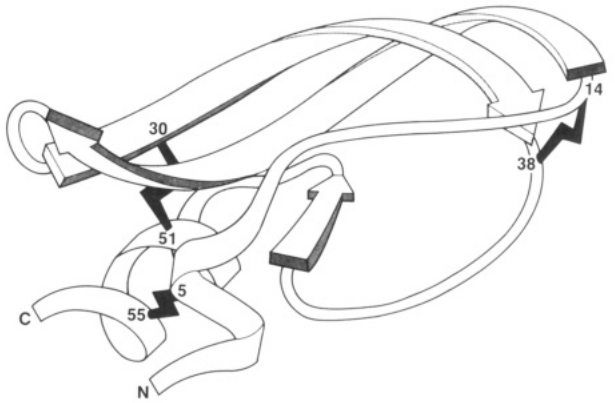


FIGURE 1: A ribbon diagram of BPTI showing the positions of the three disulfides, modified from Richardson (1981). Disulfides are made by cysteines 14 and 38, cysteines 30 and 51, and cysteines 5 and 55.

sulfide formation (Creighton, 1977a,b; Creighton & Goldenberg, 1984; Weissman & Kim, 1991). Scheme II shows a simplified version of the preferred BPTI folding mechanism, characterized at pH 8.7, 25 °C.

R is fully reduced protein containing six cysteines. Species I and II are mixtures of one- and two-disulfide bond intermediates, respectively. (14–38, 5–55) and (30–51, 5–55) are two-disulfide bond intermediates with native-like conformations. Native BPTI is represented as (14–38, 30–51, 5–55). Initial formation of a disulfide from R is most likely a nearly random event with rates of disulfide formation influenced primarily by the number of residues separating two cysteines. However, rapid intramolecular thiol-disulfide exchange among the one-disulfide bond intermediates and stabilizing structural elements present in (30–51), but not in many of the other one-disulfide bond intermediates, cause (30–51) to dominate the composition of I (Creighton, 1977c, 1988).

Continued productive folding involves (30–51). Three two-disulfide bond intermediates, (30–51, 14–38), (5–14, 30–51), and (5–38, 30–51), grouped for kinetic purposes as II, are formed directly and readily from (30–51). For productive folding to continue, thiol–disulfide rearrangements must occur in (5–38, 30–51) and (5–14, 30–51) producing (30–51, 5–55). These rearrangements are the rate-limiting steps in folding. The final step in folding is formation of the solvent-accessible 14–38 disulfide.

In addition to productive mechanisms generating a folded protein with all three disulfides, a metastable native-like intermediate missing the 30–51 disulfide, (14–38, 5–55),[3] accumulates to high levels during BPTI folding (Creighton & Goldenberg, 1984; States et al., 1984). For this intermediate, cysteines 30 and 51 are sequestered in the protein interior at the interface of the C-terminal helix and the central β-sheet. This intermediate arises through thiol–disulfide rearrangements involving other two-disulfide bond intermediates and de novo from the pool of one-disulfide bond intermediates. Since (5–55) has recently been identified as an intermediate of BPTI folding (Weissman & Kim, 1991), it is the logical direct precursor to (14–38, 5–55).

The BPTI gene has been cloned into a variety of genetically-engineered expression systems (Altman et al., 1991, and references within). A more detailed picture of BPTI folding and unfolding mechanisms should result from studies using mutants with amino acid substitutions at sites other than those of the cysteines. This approach identifies amino acid residues key to individual folding and unfolding steps. Such studies have begun (Goldenberg et al., 1989). A complementary strategy employs removal of one or more cysteines. This maneuver eliminates specific intermediates, simplifies the pathway, and clarifies the roles of certain intermediates. Exploration of energetically less favorable pathways becomes possible (Marks et al., 1987a; Goldenberg, 1988). Further, as BPTI is extremely resistant to denaturation, folding and unfolding mechanisms cannot be studied if all disulfides are present. The stability of BPTI is substantially decreased when any one disulfide is removed (Vincent et al., 1971; Schwarz et al., 1987; States et al., 1987; Hurle et al., 1990). With disulfide mutants, it is possible to compare folding mechanisms when disulfides are present (Hurle et al., 1990) and when folding is coupled to disulfide formation.

Goldenberg (1988) reported a detailed kinetic analysis of the folding and unfolding of a BPTI mutant in which cysteines 14 and 38 were replaced by serines (Ser14/Ser38). Here we report a similar analysis for a BPTI mutant in which cysteines 30 and 51 are replaced by alanines (Ala30/Ala51) and for a chemically-modified derivative of this mutant in which only the 5–55 disulfide is present (Cam14/Cam38; Ala30/Ala51). The tertiary structure of Ala30/Ala51 is nearly identical to that of BPTI (Eigenbrot et al., 1990; Hurle et al., 1991). The disulfide bond pattern is that of the native-like two-disulfide bond intermediate, (14–38, 5–55), identified as a kinetic trap in the BPTI folding mechanism (Creighton & Goldenberg, 1984; States et al., 1984). To the extent that the structural characteristics of Ala30/Ala51 are analogous to those of (14–38, 5–55), delineation of Ala30/Ala51 folding and unfolding mechanisms clarifies the mechanism of (14–38, 5–55) formation during BPTI folding. While the initial impetus of this study was further characterization of the BPTI folding pathway with Ala30/Ala51 as the simplifying model, this mutant protein proves to be an interesting model system for folding in its own right.

Materials and Methods

Materials

Protein Preparation and Purification. Plasmid constructs and expression systems for genes encoding the BPTI mutants Cys30/Cys51 → Ala30/Ala51 and Cys14/Cys38 → Ala14/Ala38 were described (Marks et al., 1987a,b). Purification and chemical characterization of these proteins were also described (Hurle et al., 1990).

The derivative (Cam14/Cam38; Ala30/Ala51) was prepared by incubating 70 μM Ala30/Ala51 with 0.5 mM DTT^{SH}_{SH} in ice-cold 0.01 M Tris-HCl, pH 8.5, 1 mM EDTA for 30 min under an argon atmosphere. Addition of an equal volume of 0.5 M iodoacetamide, 0.1 M Tris-HCl, pH 8.5, 1 mM EDTA quenched disulfide reduction. After 15 min, salts were removed by chromatography on a column (6 × 34 cm) of Sephadex G-25 equilibrated with 0.05 M ammonium bicarbonate. After lyophilization, (Cam14/Cam38; Ala30/Ala51) was separated from Ala30/Ala51 by FPLC reverse-phase PepRPC HR 10/16 chromatography using 0.1% TFA in water (A) and 0.1% TFA in acetonitrile (B). The gradient was 28–30.5% B at a flow rate of 4 mL/min for 39 min.

Tetra(carbamoylmethyl)-Ala30/51 was prepared by incubation of 3 mg of Ala30/Ala51 with a 10-fold molar excess of DTT^{SH}_{SH} for 30 min in 6 M Gdn-HCl, 0.2 M Tris, pH 8.7, 10 mM EDTA, followed by alkylation with a 20-fold molar excess of iodoacetamide over DTT^{SH}_{SH} for 20 min. Salts were removed by chromatography over Sephadex G-25 equilibrated with 10 mM HCl, and the protein solution was lyophilized.

[3] This intermediate is also known as N(30SH, 51SH) (Creighton & Goldenberg, 1984) and as N* (Staley & Kim, 1990).

Chemicals. All redox reagents were obtained from Sigma Chemical Co. and used without further purification. Solutions of DTT^{SH}_{SH} and GSH prepared by dry weight measurements contained the expected concentration of thiols when assayed by the method of Ellman (1959).

Intermolecular disulfide-containing compounds are significantly better oxidizing reagents than DTT^{S}_{S} and, if present as contaminants, alter the appearance of DTT^{S}_{S}-coupled folding profiles (Creighton, 1974a, 1977b). Crystalline DTT^{S}_{S} was assayed for such impurities using the method of Zahler and Cleland (1968). Impurities, at a level of 0.05% or greater, would have been detected. None were found. We also examined the electrophoretic profiles of reduced Ala14/Ala38 which had been incubated with 0.04 or 0.06 M DTT^{S}_{S} for 1 h. (Complete folding conditions are described below.) As expected, for DTT^{S}_{S} preparations containing insignificant amounts of impurities, no folded Ala14/Ala38 appeared (Goldenberg, 1988). Finally, we found that 1.7–2.6 μM reoxidized (Cam14/Cam38; Ala30/Ala51) accumulated within 40 min in the presence of 0.06 M DTT^{S}_{S} with no additional oxidation occurring during the next 20 min. Using the second-order rate constant for 5–55 disulfide formation by DTT^{S}_{S} (determined by other means—see results) and an initial concentration of 0.06 M DTT^{S}_{S}, simulation of reduced (Cam14/Cam38; Ala30/Ala51) folding kinetics suggested that no more than 0.3 μM oxidized (Cam14/Cam38; Ala30/Ala51) should be present after 40 min. Comparison of the experimental results and the simulation suggests that disulfides, which are less stable than DTT^{S}_{S}, were present at a level of 0.005% or less. (Residual air-oxidation might also contribute to the low levels of (Cam14/Cam38; Ala30/Ala51) oxidation in the presence DTT^{S}_{S}.) Given the results of the three different experiments described above, it was concluded that insignificant amounts of intermolecular disulfide impurities were present; consequently, DTT^{S}_{S}, was not further purified.

Methods

Identification of the Disulfide Bond in (Cam14/Cam38; Ala30/Ala51). Tetra(carbamoylmethyl)-Ala30/Ala51 and (Cam14/Cam38; Ala30/Ala51), repurified by reverse-phase FPLC chromatography after use in folding experiments, were each digested with endoproteinase Lys C (Calbiochem; EC 3.4.99.30) in 0.1 M sodium phosphate, pH 7.5, 2 M urea for 16 h at 37 °C. Protein concentrations were approximately 1 mg/mL, and enzyme to substrate ratios were approximately 0.03 unit:1 mg. Peptide fragments were isolated by FPLC reverse-phase PepRPC HR 5/5 chromatography using 0.1% TFA in water (A) and 0.09% TFA in acetonitrile (B). The gradient was 0–40% B at a flow rate of 0.4 mL/min for 110 min. Prior to tandem mass spectrometry experiments, contaminating sodium ions were removed by HPLC reverse-phase chromatography on a Vydac C-18 analytical column using 0.1% TFA in water (A) and 0.08% TFA in acetonitrile (B). The gradient was 5–50% B for 45 min at a flow rate of 1 mL/min.

Molecular masses of endoproteinase Lys C-generated peptides were determined by mass spectrometry using a Kratos MS-50S double-focusing mass spectrometer equipped with a high field magnet (mass range 3000 Da at 8 kV), a cesium ion liquid secondary ion mass spectrometry source (Aberth et al., 1982; Falick et al., 1986a), and a cooled sample introduction probe (Falick et al., 1986b). The matrix was a 1:1 mixture of glycerol and thioglycerol which was 0.1 M in HCl. Tandem mass spectrometry experiments were performed on a Kratos Analytical (Manchester, U.K.) Concept IIHH four-sector EBEB tandem mass spectrometer (Walls et al., 1990) fitted with an electro-optical array detector (Cottrell & Evans, 1987), a cesium ion liquid secondary ion source, and a coolable probe. The same matrix was used, and parent ions were generated with an 18-keV cesium ion primary beam. The collision energy for collision-induced dissociation was 4 keV. The collision gas (He) was used at a pressure sufficient to suppress the parent ion beam to about 30% of its initial value. The instrument was controlled and data were acquired with a DS-90 data system. Analysis and display were carried out with a Mach 3 data system.

Folding and Unfolding Experiments. Procedures used for folding and unfolding kinetic studies were those of Creighton and Goldenberg (1984) with modifications. These modifications do not alter folding or unfolding conditions; therefore, our kinetic measurements can be directly compared to previous work.

Protein concentrations were determined by ultraviolet spectroscopy using an extinction coefficient of 5720 cm^{-1} M^{-1} at the wavelength of maximum absorbance in the near-ultraviolet region (Kosen et al., 1981). Solutions of DTT^{S}_{S} were filtered through Gilman Acrodisc assemblies. Concentrations of DTT^{S}_{S} solutions were determined using an extinction coefficient of 110 cm^{-1} M^{-1} at 310 nm (Creighton, 1975a). Concentrations of DTT^{SH}_{SH}, GSH, and GSSG were based on dry weight measurements.

Solutions of 10 mM HCl and 1.0 M Tris-HCl, pH 9.1 (at 25 °C), 2 M KCl, 10 mM EDTA were degassed and saturated with argon. Protein and redox reagents were dissolved separately in 10 mM HCl and mixed with (one-tenth final volume) Tris/KCl/EDTA buffer plus sufficient HCl so that final solution conditions were 0.1 M Tris-HCl, pH 8.7 (at 25 °C), 0.2 M KCl, 1 mM EDTA, and 30 μM protein. Reactions of greater than 10-min duration were carried out under an argon atmosphere in septum-capped tubes.

Folding and unfolding reactions were quenched by addition of an aliquant of the reaction mixture to 0.91 M iodoacetic acid, 0.45 M Tris (free base), 0.91 M KOH, giving a final iodoacetate concentration of 0.18 M, pH ~8.5 (at room temperature). After being mixed, quenched solutions were kept at room temperature for 2 min and then placed on ice. A solution of 0.4% methyl green, 25% glycerol, 75% electrophoresis buffer was added to the quenched solutions. The final glycerol concentration was 3.4%. Samples (40–50 μL) were electrophoresed through 15% nondenaturing polyacrylamide gels (Reisfeld et al., 1962) which contained piperazine diacylamide in place of *N,N'*-methylenebisacrylamide. The electrophoresis apparatus's glass plates were silanized with Sigmacote (Sigma Chemical Co.) to prevent the gels from adhering to the glass. Electrophoresis was performed at 2–4 °C for 2.5–3 h at 80 V, followed by 4 h at 300 V.

Gels were stained and destained as described previously by Creighton and Goldenberg (1984). After being destained, dried gels were scanned with a Hoefer scanning densitometer interfaced to a Macintosh II computer (Apple Computer Corp.) running the Hoefer GS-370 data system software. Relative areas of stained protein bands were determined using the software's manual integration mode.

Kinetics of folding and unfolding of Ala30/Ala51 were simulated using the software package Stella, running on a Macintosh II. Fourth-order Runge–Kutta numerical integration was performed with a time step of 0.1. Decreasing the time step of 0.01 did not alter the simulation results. Rate constants were adjusted until simulations matched observed time dependencies for the relative amounts of the kinetic species.

Table I: Summary of Control Alkylation Experiments

alkylation conditions	% of (14–38, 5–55)	% of (5–55)	% of I	% of R
0.2 M iodoacetate	48	8	15	28
0.4 M iodoacetate	46	9	21	24
1.0 M iodoacetate	52	8	16	24
0.18 M iodoacetate[a]	48	10	17	25
0.18 M iodoacetate[b]	39 ± 2	6.8 ± 0.8	23 ± 3	31 ± 2

[a] Values obtained from the control experiment. [b] Values obtained from kinetic experiments following the reduction and unfolding of Ala30/Ala51 in the presence of 0.1 mM DTT^{SH}_{SH}. The 5-min time point is reported. Uncertainties in the measurements are reported as standard deviations.

Control Alkylation Experiments. To determine if standard alkylation conditions accurately trapped the kinetic species accumulating during Ala30/Ala51 folding and unfolding, the following experiment was performed. Four samples of Ala30/Ala51 were reduced with 0.1 mM DTT^{SH}_{SH} for 5 min. Then, the usual quench solution was added to one sample, and the following were added to the other three: one-fourth volume of 1 M iodoacetate, 0.5 M Tris-HCl, pH 8.7; one-fourth volume of 2 M iodoacetate, 0.5 M Tris-HCl, pH 8.7; or one-half volume of 2 M iodoacetate, 0.2 M Tris-HCl, pH 8.7. (The pH of the latter three quench solutions was adjusted at room temperature.) After 2 min at room temperature for the first two samples and 1 min at room temperature for the second two, the solutions were frozen until salts could be removed using a FPLC G-25 fast desalting column equilibrated with 10 mM HCl. After lyophilization, a portion of each protein mixture was separated by polyacrylamide gel electrophoresis and the relative amounts of the kinetic species were determined by densitometry. Table I summarizes the experimental results. No systematic differences in the relative amounts of the kinetic species were noted. Differences between control experiments and the average of the kinetic experiments may be ascribed to high concentrations of salts present during electrophoresis which somewhat distort the protein bands' shapes and thus decrease the accuracy of integration (in kinetic experiments) or to selective loss of unfolded protein on desalting (in control experiments).

Results

Characterization of the Disulfide Bond in (Cam14/Cam38; Ala30/Ala51). Endoproteinase Lys C peptide maps of tetra(carbamoylmethyl)-Ala30/Ala51 and (Cam14/Cam38; Ala30/Ala51) are shown in Figure 2. Using molecular masses determined by cesium ion liquid secondary ion mass spectroscopy, the labeled peptides in the chromatogram of tetra(carbamoylmethyl)-Ala30/Ala51 (Figure 2; lower tracing) were identified as the five unique fragments expected of endoproteinase Lys C proteolysis. These fragments are peptide 1 (residues 42–46), peptide 2 (residues 47–58), peptide 3 (residues 16–26 and 27–41; two peptide fragments), and peptide 4 (residues 1–15). Peptides 2 and 4, containing residues 47–58 and 1–15, respectively, were missing from the peptide map of (Cam14/Cam38; Ala30/Ala51). A new peptide appeared in that map, labeled as 4a (Figure 2; upper tracing). Peptides 1 and 3 of the tetra(carbamoylmethyl)-Ala30/Ala51 and (Cam14/Cam38; Ala30/Ala51) peptide maps were shown to be identical by mass spectrometry. The monoisotopic mass of peptide 4a (2944.4 Da) was that expected for residues 1–15 and 47–58 joined by a disulfide. To determine whether cysteine 5 or cysteine 14 formed the disulfide with cysteine 55, peptide 4a was incubated in 10 μL of 20 mM DTT^{SH}_{SH}, 0.01% NH_3(aq) for 5 min. Then, an aliquant of this mixture was added to the matrix on the probe of the tandem mass spectrometer. A peak was present in the normal mass spectrum of the mixture at 1779.8 Da (monoisotopic mass) corresponding to the protonated peptide consisting of residues 1–15. This peak was selected in the first mass spectrometer and subjected to collision-induced dissociation with He gas. The fragment ions were analyzed in the second mass spectrometer. As seen in Table II, the collision-induced dissociation spectrum is entirely consistent with a cysteine at position 5 and a (carbamoylmethyl)cysteine at position 14. Therefore, the single disulfide in (Cam14/Cam38; Ala30/Ala51) links cysteines 5 and 55.

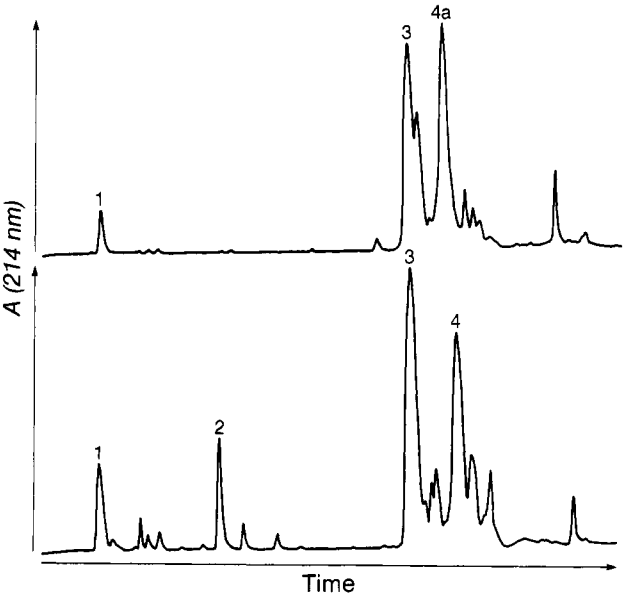


FIGURE 2: Reverse-phase chromatography of endoproteinase Lys C peptides of tetra(carbamoylmethyl)-Ala30/Ala51 (lower tracing) and (Cam14/Cam38; Ala30/Ala51) (upper tracing). Identities of the labeled peaks are given in the text. The unlabeled peak following that labeled 3 in the upper tracing was shown to be residues 27–46 by mass spectrometry.

Reduction of the 5–55 Disulfide in (Cam14/Cam38; Ala30/Ala51) by DTT^{SH}_{SH}. Figure 3 shows the electrophoretic pattern of reduced and oxidized (Cam14/Cam38; Ala30/

Table II: Observed Sequence Ions in the High-Energy Collision-Induced Dissociation Mass Spectrum of Residues 1–15 (MH^+ = 1779.8 Da) Derived from the Disulfide-Linked Fragments 1–15 and 47–58[a]

a			341	488	591	704	833	930		1190	1291	1348	1445	1605	
b		254	369		619	732	861		1055	1218	1319			1633	
c				533		749				1235	1336		1490	1650	
d					559									1516	
	Arg	Pro	Asp	Phe	Cys	Leu	Glu	Pro	Pro	Tyr	Thr	Gly	Pro	Cam	Lys
w						1102	973		779						
x				1437		1187									
y	1179		1526	1411				919		725					
z						1145									

[a] Fragment ion masses in the table are nominal monoisotopic masses in daltons. The data confirm unequivocally that residue 5 is Cys and residue 14 is Cam. The nomenclature used follows that of Biemann (1990). Cam = (carbamoylmethyl)cysteine.

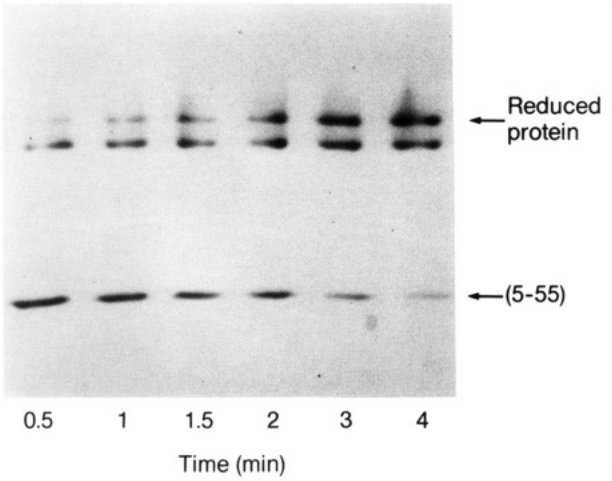


FIGURE 3: Electrophoretic pattern of species trapped by alkylation with iodoacetate during the reduction of the 5–55 disulfide of (Cam14/Cam38; Ala30/Ala51) by 10 mM DTT^{SH}_{SH} at 25 °C, pH 8.7. The initial concentration of (Cam14/Cam38; Ala30/Ala51) was approximately 27 μM. Quench times are indicated in the figure. Positions of (Cam14/Cam38; Ala30/Ala51), as (5–55), and reduced protein blocked with carbamoylmethyl groups at cysteines 14 and 38 and carboxymethyl groups at cysteines 5 and 55 are indicated in the figure. The unlabeled band, migrating more rapidly than reduced protein, is contaminating protein containing carbamoylmethyl groups at all four cysteines. This material comigrated with (Cam14/Cam38; Ala30/Ala51) during reverse-phase chromatography; as it does not interfere with kinetic analysis, further purification was not attempted.

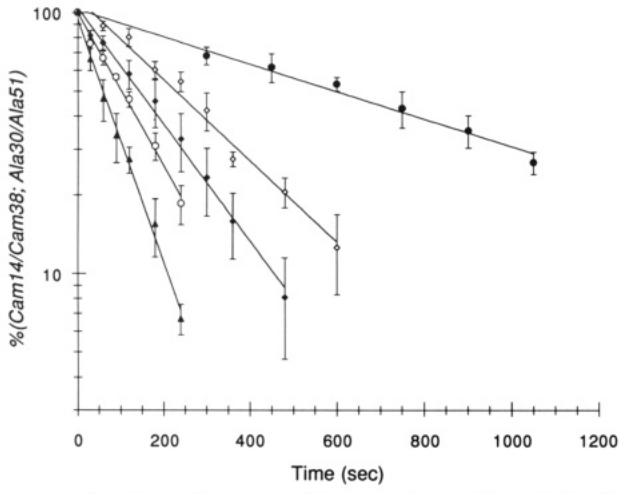


FIGURE 4: Semi-logarithmic plot of the percentage of (Cam14/Cam38; Ala30/Ala51) present versus time after addition of DTT^{SH}_{SH}, pH 8.7, 25 °C. The concentrations of DTT^{SH}_{SH} used were 1 mM (●), 2 mM (◇), 4 mM (◆), 6 mM (○), and 10 mM (▲). Error bars are standard deviations of experimental values. Solid lines are linear least-squares fits of the data, and slopes of these lines are pseudo-first-order rate constants.

Ala51) trapped by carboxymethylation after incubation times were increased with an initial concentration of 10 mM DTT^{SH}_{SH}. The rate of reduction of the 5–55 disulfide in the presence of 10 mM DTT^{SH}_{SH} followed pseudo-first-order kinetics, and similar kinetics were observed over a 10-fold DTT^{SH}_{SH} concentration range (Figure 4). When pseudo-first-order rate constants, obtained from slopes of the curves in Figure 4, are plotted against DTT^{SH}_{SH} concentration, the relationship between the two variables is a straight line with a slope of 1.0 s^{-1} M^{-1}. This value is the second-order rate constant (k_{-1}) for reduction of the 5–55 disulfide in (Cam14/Cam38; Ala30/Ala51) by DTT^{SH}_{SH}.

Oxidation of the 5–55 Disulfide in (Cam14/Cam38; Ala30/Ala51) in a Mixed Glutathione Redox Environment. Scheme III

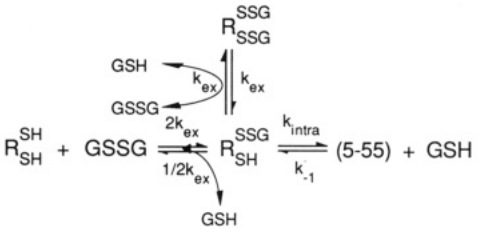


Table III: Intramolecular Rate Constants for Folding of (Cam14/Cam38; Ala30/Ala51) in the Presence of Oxidized and Reduced Glutathione

[GSH] (M)	[GSSG] (M)	k_{intra}[a] (s^{-1})	k_{intra}[b] (s^{-1})
0.01	0.0001	0.023	0.033
0.01	0.005	0.0073	0.029
0.01	0.01	0.0027	0.024
0.015	0.001	0.024	0.030
0.02	0.0005	0.023	0.028
0.0005	0.001	0.014	0.026

[a] Calculated using eq 1. [b] Calculated using eq 2.

Formation of the 5–55 disulfide by DTT^{S}_{S} occurs extremely slowly. It is therefore difficult to measure accurately the rate of 5–55 disulfide formation in the presence of DTT^{S}_{S}. GSSG is a more powerful oxidizing reagent, with a greater redox potential, than DTT^{S}_{S}. It will promote disulfide formation in cases where reaction with DTT^{S}_{S} occurs slowly or not at all (Creighton & Goldenberg, 1984; Goldenberg, 1988). The mechanism of formation of the 5–55 disulfide in the presence of GSSG is expected to follow Scheme III.

The rate constants k_{intra} and k_{-1} have been defined above. The value for k_{-1} is assumed to be identical for reduction of the 5–55 disulfide by DTT^{SH}_{SH} and GSH (Creighton & Goldenberg, 1984). The rate constant k_{ex} reflects an average rate of formation of the mixed disulfide between oxidized glutathione and either one of the two protein thiols. Values of 2 and $^1/_2$ are statistical correction factors for formation and reduction of the initial mixed disulfide. Scheme III includes the possibility that R^{SSG}_{SSG}, with mixed disulfides at both cysteines 5 and 55, and/or R^{SSG}_{SH} with a mixed disulfide at either cysteine 5 or 55 accumulate(s). A value for k_{intra} can be determined by following the kinetics of folding of reduced (Cam14/Cam38; Ala30/Ala51) in the presence of a mixed redox buffer containing GSSG and GSH (Goldenberg, 1988), by determining the third-order rate constant for reduction of the 5–55 disulfide by GSH (Creighton & Goldenberg, 1984) or by determining the relative amounts of reduced and oxidized protein present at equilibrium.

We used the equilibrium experiment to determine k_{intra}. Reduced (Cam14/Cam38; Ala30/Ala51) was incubated with six different mixtures of GSSG and GSH (Table III) for 2 h under standard folding conditions. Then, iodoacetate was added and a portion of the mixture was subjected to electrophoresis. Fractions of reduced ($f_{R_{total}}$) and oxidized ($f_{(5-55)}$) (Cam14/Cam38; Ala30/51) were determined by densitometry.

For Scheme III, if $k_{intra} \gg 2k_{ex}$[GSSG], then at equilibrium, k_{intra} will be related to k_{-1}, [GSSG], [GSH], $f_{R_{total}}$, and $f_{(5-55)}$ as

$$k_{intra} = \frac{f_{(5-55)}k_{-1}[GSH]^2}{f_{R_{total}} \times 4[GSSG]} \quad (1)$$

On the other hand, if $k_{intra} \ll 2k_{ex}$[GSSG] then

$$k_{intra} = \frac{f_{(5-55)}k_{-1}([GSH] + 2[GSSG])^2}{f_{R_{total}} \times 4[GSSG]} \quad (2)$$

Table IV: Apparent Rate Constants for the Folding and Unfolding Transitions of Ala30/Ala51 in the Presence of DTT^{S}_{S} and DTT^{SH}_{SH}[a]

A. Steps Involving Intermolecular Thiol–Disulfide Exchange

step	formation (s^{-1} M^{-1})	breakage (s^{-1} M^{-1})	intramolecular[b] (s^{-1})
$R \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} I$	0.019 (0.59; 0.89)[c]	72 (1.9; 3.6)[c]	5.8
$(5–55) \underset{k_{-4}}{\overset{k_4}{\rightleftharpoons}} (14–38, 5–55)$	11.5 (1.2; 2.0)[d]	43 (1.1; 1.4)[d]	3.45×10^3
$R \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} (5–55)$	9.3×10^{-5} (17.5)[e]	1.0 (8300)[e]	0.028

B. Step Involving Intramolecular Thiol–Disulfide Exchange

step	forward (s^{-1})	reverse (s^{-1})
$I \underset{k_{-3}}{\overset{k_3}{\rightleftharpoons}} (5–55)$	0.03 (6)[f]	0.09 (4500; 7500)[f]

[a] Values in parentheses are the ratios of rate constants for folding or unfolding of the Ala30/Ala51 species compared to those of BPTI or Ser14/Ser38 at the corresponding kinetic step. [b] $k_{obs} = k_{intra} \times 3.3 \times 10^{-3}$ M^{-1} (Creighton & Goldenberg, 1984). [c] Comparison is made with the disappearance or appearance of reduced protein in the BPTI folding pathway (Creighton, 1977b; Creighton & Goldenberg, 1984). [d] Comparison is made with formation or reduction of the 14–38 disulfide in the BPTI folding pathway (Creighton & Goldenberg, 1984; Creighton, 1975a). [e] Comparison is made with direct appearance or disappearance of (5–55, 30–51) from or to (30–51) in the folding and unfolding mechanisms of Ser14/Ser38 (Goldenberg, 1988). [f] Comparison is made with the appearance or disappearance of (30–51, 5–55) by intramolecular thiol–disulfide exchange in the BPTI folding pathway (Creighton & Goldenberg, 1984; Goldenberg, 1988).

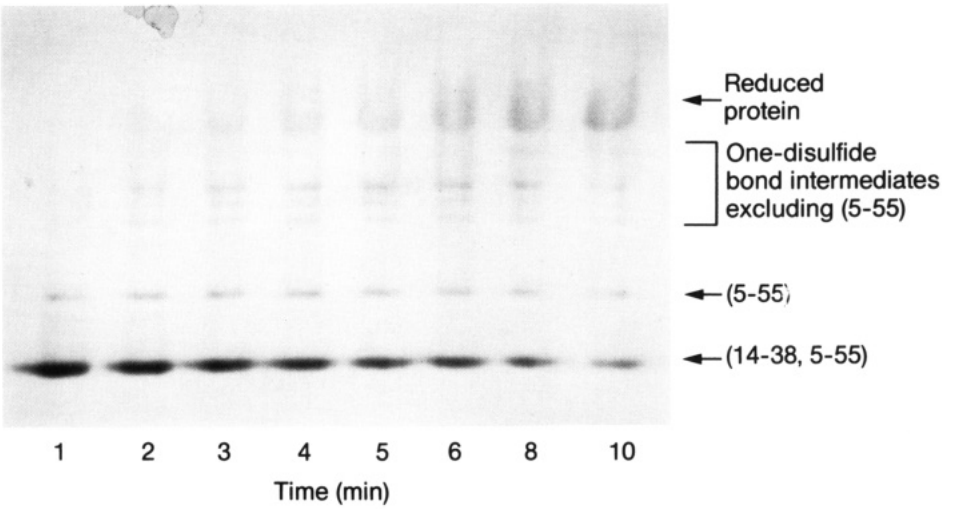


FIGURE 5: Electrophoretic pattern of species trapped by alkylation with iodoacetate during unfolding of Ala30/Ala51 in the presence of 0.1 mM DTT^{SH}_{SH} at 25 °C, pH 8.7. Quench times are indicated in the figure. The various kinetic species are given by labels in the figure.

For the first case, mixed disulfides between glutathione and reduced protein should not accumulate and only fully reduced protein (R^{SH}_{SH}) contributes to the value of $f_{R_{total}}$. For the second case, $f_{R_{total}}$ may include contributions by R^{SH}_{SH}, R^{SSG}_{SH}, and R^{SSG}_{SSG} depending on the ratio of [GSSG] to [GSH]. Both a glutathione–protein mixed disulfide and carboxymethylation of a cysteine introduce one negative charge per modified cysteine so that after alkylation R^{SH}_{SH}, R^{SSG}_{SH}, and R^{SSG}_{SSG} are indistinguishable on gel electrophoresis; but as $f_{R_{total}}$ incorporates all three compounds, it is not necessary to distinguish among the various unfolded species. Large excesses of GSH and GSSG were used relative to protein; therefore, corrections were not made to the equilibrium redox potential. Consistent values for k_{intra} were calculated using eq 2, but not eq 1. A summary of these experiments is presented in Table III. The average value for k_{intra} is 0.028 ± 0.004. Using eq 3, given below, this value can be converted to an apparent second-order rate constant (k_1) for formation of the 5–55 disulfide by DTT^{S}_{S} (Table IV). Knowing k_1 and k_{-1}, we can evaluate the contribution of direct 5–55 disulfide formation and reduction to the more complex folding mechanism of Ala30/Ala51.

Unfolding and Folding of Ala30/Ala51 in the Presence of DTT^{SH}_{SH} and DTT^{S}_{S}. Figure 5 shows the electrophoretic pattern of kinetic species trapped by iodoacetate during Ala30/Ala51 unfolding in the presence of 0.1 mM DTT^{SH}_{SH}, pH 8.7, 25 °C. Additional unfolding studies were performed with 0.2 or 0.3 mM DTT^{SH}_{SH} and in the presence of a mixture of 0.3 mM DTT^{SH}_{SH} and 0.01 M DTT^{S}_{S}.

Of the seven protein bands, the identities of reduced protein and Ala30/Ala51 were confirmed by comparison of migration rates with authenticated samples. A protein isolated using (carboxymethyl)cellulose and reverse-phase chromatography and shown by endoproteinase Lys C peptide mapping and tandem mass spectrometry to contain only the 5–55 disulfide has a migration rate identical to that labeled (5–55) in Figure 5. (Data not shown.) The other four bands were assumed, on the basis of migration patterns, to be one-disulfide bond intermediates. All one-disulfide bond intermediates after carboxymethylation have a nominal charge difference of −2 compared to Ala30/Ala51 and +2 compared to fully reduced and alkylated protein. One-disulfide bond intermediates should migrate between fully reduced and fully oxidized protein due to charge differences. Additional differences in migration rates can be attributed to differences in hydrodynamic volumes. The intermediate, (5–55), migrates as if it has a compact shape, whereas the other one-disulfide bond intermediates migrate as if they are much more disordered.

For a protein with four cysteines, six one-disulfide bond

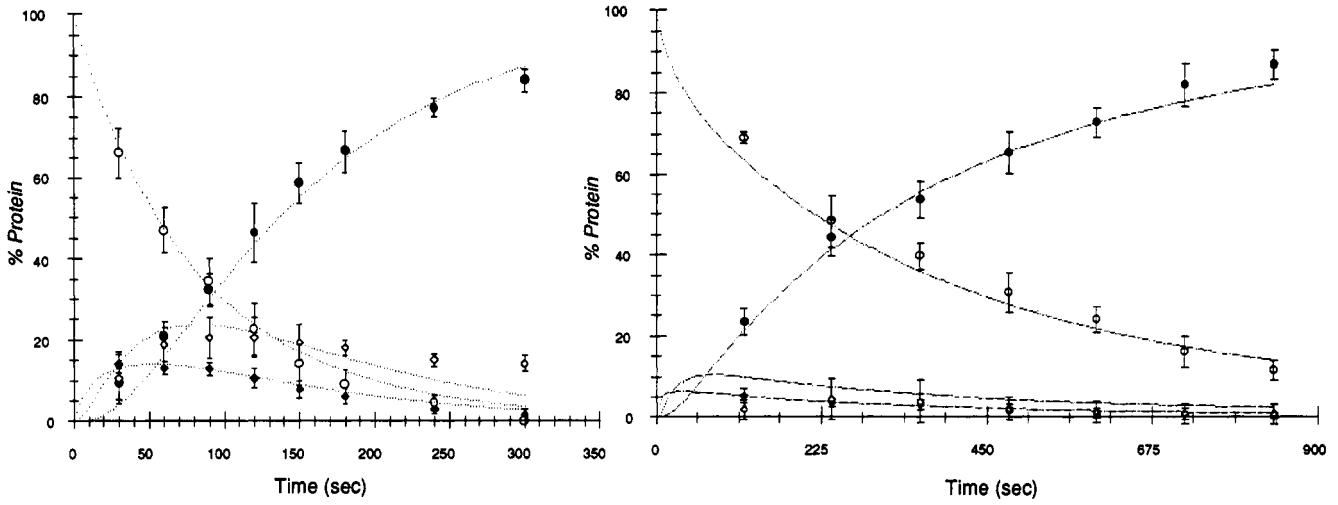


FIGURE 6: Kinetics of unfolding of Ala30/Ala51 in the presence of 0.3 mM DTT^{SH}_{SH} (left-hand panel) or 0.01 M DTT^{S}_{S} and 0.3 mM DTT^{SH}_{SH} (right-hand panel). Symbols show the relative amounts of reduced protein (●), one-disulfide bond intermediates excluding (5–55) (◇), (5–55) (◆), and Ala30/Ala51 (○). Dotted lines are the results of simulations using Scheme IV and rate constants in Table IV. For simulations, points for every two time steps are plotted. Other experimental details are given in the text. Error bars are standard deviations of experimental values.

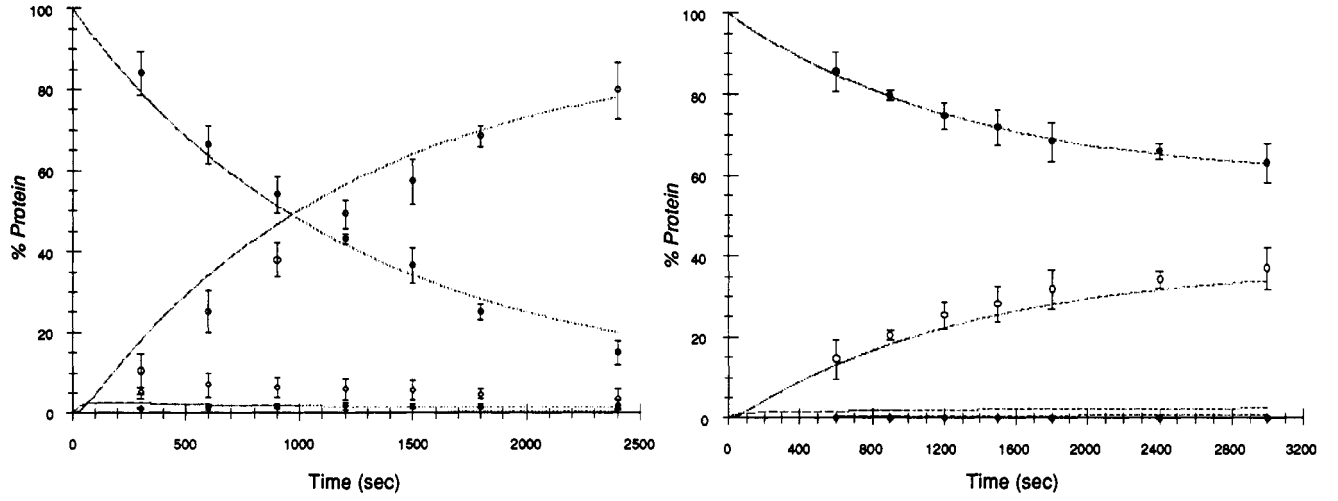


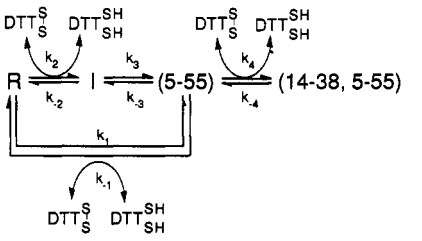
FIGURE 7: Kinetics of folding of Ala30/Ala51 in the presence of 0.04 M DTT^{S}_{S} (left-hand panel) or 0.02 M DTT^{S}_{S} and 0.1 mM DTT^{SH}_{SH} (right-hand panel). Symbols have the same meaning as in Figure 6. Dotted lines are simulation results, as in Figure 6. Error bars are standard deviations of experimental values. Experimental details are given in the text.

intermediates are possible. [Mixed disulfides between DTT and protein thiols are not expected (Creighton & Goldenberg, 1984).] Four one-disulfide bond intermediates might be formed by direct thiol–disulfide rearrangement with (5–55) as the reactant. The other one-disulfide bond intermediate, (14–38), cannot be formed directly from (5–55). Formation of (14–38) from (5–55) requires two intramolecular rearrangements. We do not know if the five gel bands, corresponding to one-disulfide intermediates, contain all six possible intermediates. Identities of one-disulfide bond intermediates other than (5–55) are now being determined.

Folding of reduced Ala30/Ala51 was followed in a manner identical to that of unfolding, except DTT^{S}_{S} (or a mixture of DTT^{SH}_{SH} and DTT^{S}_{S}) was present. The following redox conditions were used: 0.01 M, 0.02 M, or 0.04 M DTT^{S}_{S} and a mixture of 0.02 M DTT^{S}_{S} and 0.1 mM DTT^{SH}_{SH}. Two one-disulfide bond intermediates accumulated to barely detectable levels during folding in the presence of DTT^{S}_{S}.

All Ala30/Ala51 folding and unfolding kinetic data were simulated using a single kinetic model (Scheme IV) with a unique rate constant assigned to each folding and unfolding step (Table IV). The one-disulfide bond intermediates, excluding (5–55), were treated as a kinetically homogeneous

Scheme IV



population. All steps except I ↔ (5–55) involve intermolecular thiol–disulfide exchange reactions with DTT^{S}_{S} or DTT^{SH}_{SH}. Interconversion of I to (5–55) and the reverse reaction involve intramolecular thiol–disulfide bond rearrangements. Simulations included direct interconversion of R and (5–55), using the apparent second-order rate constants k_1 and k_{-1}, although contributions by this step to the observed appearance and or disappearance of Ala30/Ala51 are negligible. Simulations and experimental data are shown in Figures 6 and 7 for representative unfolding and folding experiments.

Oxidized DTT is a particularly useful disulfide reagent since the observed second-order rate constant for protein disulfide formation during folding is directly related to the first-order

rate constant of polypeptide chain collision (k_{intra}, Scheme I; Creighton, 1977b; Creighton & Goldenberg, 1984). The oxidation mechanism follows the general outline of Scheme I, but a thiol compound is not released on forming the initial mixed disulfide. A rapid-equilibrium assumption is postulated for the first step in Scheme I because reformation of the intramolecular DTT^S_S disulfide usually occurs much more rapidly than does attack of the DTT–protein mixed disulfide by a second protein thiol. The observed second-order rate constant for disulfide formation ($k_{obs} = k_1$, k_2, k_3, or k_4 in Scheme IV) is related to k_{intra}, an equilibrium constant for formation of the mixed disulfide (K) and a statistical factor of 2 (two DTT–protein mixed disulfides can be formed) as

$$k_{obs} = 2Kk_{intra} \quad (3)$$

Creighton and Goldenberg (1984) determined that $K = 1.67 \times 10^{-3}$ M^{-1} at pH 8.7, 25 °C. Using this equilibrium constant, apparent intramolecular rate constants were calculated and are given in Table IV. The apparent second-order rate constant, k_1, for direct formation of the 5–55 disulfide is also given in that table. All rate constants are underestimates since they are not corrected for partial ionization of the cysteines. Even with this caveat, comparisons of the Ala30/Ala51 folding mechanism with those of BPTI and Ser14/Ser38 (Creighton & Goldenberg, 1984; Goldenberg, 1988) are possible because environmental conditions were the same. Values in parentheses in Table IV relate the rate constants reported here to those determined previously for the BPTI and Ser14/Ser38 folding pathway.

Discussion

The experiments presented in this study directly show that the DTT^{SH}_{SH}-coupled unfolding mechanism for Ala30/Ala51 is an indirect pathway in which "correct" disulfides are *not* reduced sequentially. The first step in unfolding is reduction of the solvent-exposed 14–38 disulfide. The next step is an intramolecular thiol–disulfide rearrangement resulting in any one or all four of the intermediates, (5–14), (5–38), (14–55), and (38–55). The intermediate (14–38) cannot be formed directly from (5–55) as two thiol–disulfide exchange reactions are required. Depending on the relative rates of rearrangement versus further reduction, (14–38) may be present in the mixture of one-disulfide bond intermediates. The final step in unfolding is reduction of one or more of the one-disulfide bond intermediates, excluding (5–55). There is no absolute kinetic block on direct reduction of the 5–55 disulfide. However, given the concentrations of DTT^{SH}_{SH} used in these experiments, the rearrangement pathway is greatly favored. Only if $[DTT^{SH}_{SH}]$ was approximately 100-fold greater than used in the experiments reported here would direct reduction of the 5–55 disulfide be competitive with the rearrangement step. The folding mechanism for Ala30/Ala51 appears to be the reverse of the unfolding mechanism. Direct formation of (5–55) from R can never be competitive with the rearrangement pathway when DTT^S_S is the oxidizing reagent. The favored Ala30/Ala51 folding and unfolding mechanisms are described by the upper portion of Scheme IV.

In the following sections, we first review and interpret the experimental observations that form the foundation for the proposed folding and unfolding mechanisms of Ala30/Ala51. Next, the kinetics and energetics of individual intramolecular steps are examined. We then consider the relationship of the Ala30/Ala51 folding and unfolding mechanisms to those of (14–38, 5–55) in the BPTI folding pathway. Effects of eliminating the 30–51 disulfide on the BPTI folding pathway and of eliminating the 14–38 disulfide on the Ser14/Ser38 folding pathway are explored. Finally, disulfide bond-coupled mechanisms are compared to denaturant-dependent folding mechanisms for BPTI mutants missing the 30–51 disulfide or the 14–38 disuflide.

Validity of the Mechanism. (*a*) *Evidence That Most One-Disulfide Bond Intermediates Interconvert Rapidly with Each Other.* We have not carried out a complete folding study using GSSG instead of DTT^S_S; a few time-course studies using 0.075 mM GSSG have been performed, and the one-disulfide bond intermediate pattern seems to be virtually the same as that found during unfolding studies using DTT^{SH}_{SH}. (Data not shown.) Therefore, the same one-disulfide bond intermediates are formed regardless of the thiol–disulfide exchange reagent used. Further, the relative amounts of one-disulfide bond intermediates, excluding (5–55), appear to be the same for the two redox reagents. This finding is consistent with the view that one-disulfide bond intermediates, again excluding (5–55), rapidly interconvert and only more slowly interconvert to (5–55). Creighton (1977c, 1988) found that many different one-disulfide bond intermediates were generated during BPTI folding when Gdn-HCl was present. This mixture of one-disulfide bond intermediates, excepting (5–55), interconverted within 5 s to the usual pattern of one-disulfide bond intermediates when it was incubated in the standard folding buffer in the absence of redox reagents. Our estimate of a half-time of about 8 s for the conversion of (5–55) to I is consistent with Creighton's finding that (5–55) does not rapidly equilibrate with other one-disulfide bond intermediates on the second time scale.

(*b*) *Evidence That a Single Mechanism Is Consistent with the Observed Kinetic Data.* The *overall* rate of oxidation of R (controlled primarily by $k_2[DTT^S_S]$) is always less than the rates of the subsequent steps, given the concentrations of DTT^S_S used. Consequently, the one-disulfide bond intermediates accumulate only to barely detectable levels and the rate constants k_3, k_{-3}, k_4, and k_{-4} must be derived from unfolding studies. However, independent measures of k_{-2} are obtained from unfolding and folding studies. The *overall* rate of reduction of I controls the appearance of R during unfolding so that all intermediates accumulate and k_{-2} is extracted directly from unfolding simulations. An independent estimate of k_{-2} was also obtained from folding studies in which a mixture of DTT^S_S and DTT^{SH}_{SH} was present since the term $k_{-2}[DTT^{SH}_{SH}]$ then influences the rate of disappearance of R. Since the relative amount of R is controlled in both folding and unfolding experiments by the same rate constant, it is likely that R is derived from the same population of one-disulfide bond intermediates.

Figure 6 (left-hand panel) shows that the one-disulfide bond intermediates, labeled I, accumulated in an unfolding experiment to their maximum level at a later time than did (5–55). This behavior, which was also obvious when 0.2 or 0.3 mM DTT^{SH}_{SH} was present, indicates that these one-disulfide bond intermediates are kinetically distinguishable from (5–55). When all one-disulfide bond intermediates were treated as a single kinetic population, it was possible to fit a consistent set of rate constants to the observed unfolding kinetics; but a different rate constant was then needed to account for reduction of one-disulfide bond intermediates to reduce protein when R was folded in a mixture of DTT^S_S and DTT^{SH}_{SH}. (Simulations not shown.) The principle of microscopic reversibility can be violated since equilibrium conditions do not apply (Krupka et al., 1966); however, postulating two different mechanism, one for folding and one for unfolding, is the less

attractive alternative—especially since the data are consistent with a reversible mechanism.

Direct formation of Ala30/Ala51 from (14–38) cannot be definitively excluded from the folding mechanism. For this reaction to be competitive with the measured rate constant for rearrangement of I to (5–55), a pseudo-first-order rate constant for direct formation of the 5–55 disulfide must be about 0.03 s^{-1} so that the second-order rate constant would be about 0.75–3 s^{-1} M^{-1}. Again, as the data are consistent with a single reversible mechanism, and there is no conclusive evidence that Ala30/Ala51 is formed directly from (14–38) in the presence of DTT^S_S, this reaction is not included in the proposed DTT^S_S-coupled folding mechanism.

(c) *Evidence That (Cam14/Cam3838; Ala30/Ala51) Is an Appropriate Model for (5–55) in the Ala30/Ala51 Folding Mechanism.* Formation of (5–55) from R occurs at a rate that is 0.5% (9.3×10^{-5} s^{-1} M^{-1}/0.019 s^{-1} M^{-1}; Table IV) the overall disappearance of R. This conclusion is based on (Cam14/Cam38; Ala30/Ala51) folding studies. These studies are open to the objection that the bulky carbamoylmethyl groups affect rates of oxidation and reduction during (Cam14/Cam38; Ala30/Ala51) folding and unfolding reactions. The same objection applied to folding studies on BPTI derivatives chemically blocked at cysteines 14 and 38 (Creighton, 1977a). In that case, the criticism was answered by comparative folding studies using BPTI mutants in which cysteines 14 and 38 were replaced with other amino acids (Marks et al., 1987a; Goldenberg, 1988). Goldenberg (1988) compared the equilibrium constant for formation of (30–51, 5–55) from the one-disulfide bond intermediates of BPTI with that of Ser14/Ser38 and its one-disulfide bond intermediates. The two equilibrium constants are very similar, which argues that interpretations of the folding kinetics of wild-type and mutant protein are correct. The rates of direct 5–55 disulfide formation and reduction were also comparable whether cysteines 14 and 38 were chemically blocked or mutated.

As equilibrium analysis can also be made for formation of (5–55) from R in the folding mechanisms of Ala30/Ala51 and (Cam14/Cam38; Ala30/Ala51). Rate constants for formation and reduction (k_1 and k_{-1}) of the 5–55 disulfide in (Cam14/Cam38; Ala30/Ala51) can be used to calculate an equilibrium constant:

$$K_{eq} = \frac{[(\text{Cam14/Cam38; Ala30/Ala51})][DTT^{SH}_{SH}]}{[R][DTT^S_S]} = \frac{k_1}{k_{-1}} \quad (4)$$

A similar calculation can be made for the equilibrium constant for formation of (5–55) from reduced Ala30/Ala51 using the rate constants k_2, k_{-2}, k_3, and k_{-3}.

$$K_{eq} = \frac{[(5\text{–}55)][DTT^{SH}_{SH}]}{[R][DTT^S_S]} = \frac{k_2 k_3}{k_{-2} k_{-3}} \quad (5)$$

The equilibrium constants calculated using eqs 4 and 5 are 9.3×10^{-5} and 9.0×10^{-5}, respectively. The agreement between these two values indicates that carbamoylmethylation of cysteines 14 and 38 does not influence the stability of the (5–55) intermediate. It is unlikely, therefore, that the measured values of k_1 and k_{-1} are experimental artifacts due to carbamoylmethyl groups.

Kinetics and Energetics of Ala30/Ala51 and (Cam14/Cam38; Ala30/Ala51) Folding Transitions. Reaction coordinate diagrams for the Ala30/Ala51 and (Cam14/Cam38; Ala30/Ala51) folding mechanisms are shown in Figure 8 for redox conditions where the ratio of DTT^S_S to DTT^{SH}_{SH} is 1000:1. Apparent free energies of the reduced proteins were arbitrarily

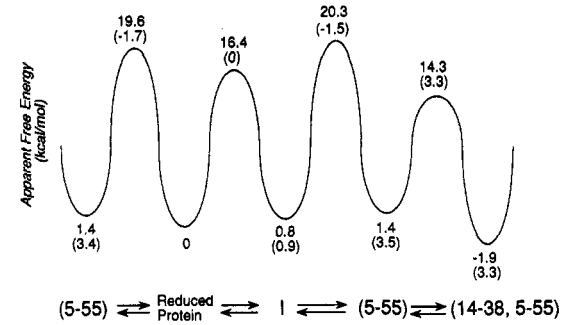


FIGURE 8: Free energy profiles for the folding mechanisms of Ala30/Ala51 and (Cam14/Cam38; Ala30/Ala51) with a DTT^S_S to DTT^{SH}_{SH} ratio of 1000:1 at 25 °C. The relative free energies of the reduced proteins were set to zero. The folding mechanism for Ala30/Ala51 is diagrammed to the right of the position for reduced protein, while the folding mechanism for (Cam14/Cam38; Ala30/Ala51) is diagrammed to the left of the position for reduced protein. For transitions involving intermolecular thiol–disulfide exchange reactions, relative stabilities of folding species were calculated according to eq 6 (given below) where rate constants k_f and k_r are observed second-order rate constants for the appropriate step, P^S_S and P^{SH}_{SH} are the oxidized and reduced protein species involved in the transition, R is the gas constant, and T is the temperature. For the transition I ↔ (5–55), the relative free energy difference was calculated using eq 7 (given below). Apparent transition state free energies were calculated using eq 8 (given below) where h is the Planck constant and k_b is the Boltzmann constant. Other terms have be defined previously. In the figure, values in parentheses give $\Delta\Delta G$ for various intermediates and transition states comparing the BPTI mechanism to the Ala30/Ala51 mechanism (Creighton, 1975a; Creighton, 1977b; Creighton & Goldenberg, 1984) or the Ser14/Ser38 mechanism to the (Cam14/Cam38; Ala30/Ala51) mechanism (Goldenberg, 1988).

$$\Delta G = -RT \ln\left(\frac{[P^S_S]}{[P^{SH}_{SH}]}\right) = -RT \ln\left(\frac{1000 k_f}{k_r}\right) \quad (6)$$

$$\Delta G = -RT \ln\left(\frac{[(5\text{–}55)]}{[I]}\right) = -RT \ln\left(\frac{k_3}{k_{-3}}\right) \quad (7)$$

$$\Delta G^{\ddagger} = RT \ln\left(\frac{k_b T}{h k_{intra}}\right) \quad (8)$$

set ot zero, a formalism used previously (Creighton, 1977b; Creighton & Goldenberg, 1984; Beasty et al., 1986, Goldenberg et al., 1989). Relative free energies of other species were then calculated using eq 6 or 7 for transitions that respectively did or did not involve intermolecular thiol–disulfide exchange. Heights of activation barriers were calculated using eq 8 (Creighton, 1977b; Creighton & Goldenberg, 1984). (See the legend of Figure 8 for eqs 6–8.) Figure 8, as a representation of the Ala30/Ala51 and (Cam14/Cam38; Ala30/Ala51) folding mechanisms, accentuates intramolecular transitions and diagrams relative stabilities of the various species *for one set of redox conditions.*

The one-disulfide bond intermediates, grouped together as I, are separated from (5–55) by the highest energy barrier in Figure 8. This barrier defines the rate-limiting step in folding and unfolding of Ala30/Ala51 for intramolecular transitions. The half-time for intramolecular unfolding of (5–55) is approximately 8 s and that for folding is about 23 s. Energy barriers for the other two steps in the Ala30/Ala51 folding pathway are considerably lower and reflect more rapid intramolecular interconversions. The half-time for formation of I from R is approximately 120 ms and that for formation of the 14–38 disulfide is approximately 200 μs.

The greater height of the activation barrier (by about 3 kcal/mol) for direct formation of the 5–55 disulfide from reduced protein compared to that for formation of the other one-disulfide bond intermediates again emphasizes the neg-

ligible contribution of direct formation of the 5–55 disulfide to the folding mechanism of Ala30/Ala51. This observation raises the following questions: what factors contribute to the slower rate of direct formation of the 5–55 disulfide (than other single disulfides) from R and what factors favor intramolecular thiol–disulfide rearrangement steps in the Ala30/Ala51 folding mechanism?

Slow formation of the 5–55 disulfide might be due to one or more of three factors: aberrant cysteine pK_a's, decreased reactivity because of steric inaccessibility either on forming DTT–protein mixed disulfides or protein–protein disulfides, or an inherently low probability of chain collision between cysteines 5 and 55. The observed rate depends on the concentration of the ionized thiols. The pK_a of cysteine 5 in a model peptide (residues 1–7 of BPTI) is approximately 0.5–0.6 pH unit higher than that of a typical protein cysteine (L. C. Ma and S. Anderson, unpublished results). The pK_a's of cysteines 14 and 38 have been estimated as 8.8 (Creighton, 1975a). We do not have a measure for the pK_a of cysteine 55. If the increased pK_a of cysteine 5 is the limiting factor, a 2–3-fold rate reduction in disulfide formation is expected. Creighton (1975b) concluded that all cysteines of reduced BPTI were approximately equally reactive to disulfide reagents. Therefore, differential formation of mixed DTT–protein disulfides is unlikely to be primarily responsible for differences in rates. Specific steric interactions preventing formation of certain protein disulfide pairs might be important. Such interactions might be important in the rate of formation of the 51–55 disulfide during BPTI folding but do not seem to contribute in a major way to formation of the first disulfide during folding of Ala30/Ala51. (See below.)

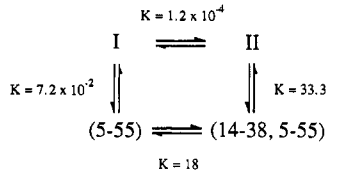
According to statistical mechanics calculations, the probability of loop formation involving two monomers in a chain is related to n^{-x}, where n is the number of virtual bonds separating two monomers in close spatial contact; x varies between $^3/_2$ (for a random-flight type chain) and about $^5/_2$ (for a chain characterized by excluded volume effects; Jacobson & Stockmayer, 1950; Chan & Dill, 1990). Thus, as a first approximation, the probability of disulfide formation decreases as the number of residues between cysteines forming the bond increases. The observed rate constant $(k_2 + k_1)$ for the disappearance of R during folding of Ala30/Ala51 is equal to the sum of the microscopic rate constants for forming each disulfide. Therefore, if the relationship between rates and chain separation is valid, we should find a correlation between the observed rate constant for forming the 5–55 disulfide (k_1) and one calculated using the equation

$$k_{\text{calcd}}^{(5-55)} = \frac{n_{(5-55)}^{-x}}{\sum_i n_i^{-x}}(k_2 + k_1) \qquad (9)$$

If x is $^3/_2$, $k_{\text{calcd}}^{(5-55)}$ is 7.7×10^{-4} s^{-1} M^{-1}. If x is $^5/_2$, $k_{\text{calcd}}^{(5-55)}$ is 2.0×10^{-4} s^{-1} M^{-1}. On the basis of statistical considerations, 5–55 disulfide formation is then expected to occur at about 1–4% of the rate observed for overall disappearance of R. These calculated rates are in reasonable agreement with the observed 200-fold rate reduction. The primary cause of the slow rate of direct formation of the 5–55 disulfide can be explained by the separation of the two cysteines along the chain, while differences in pK_a values or steric effects may be responsible for the 2–8-fold discrepancy between calculated and observed rate constants. Measurement of rates of formation of individual disulfides may be useful gauges of conformation in unfolded proteins.

The apparent intramolecular rate constants for forming the 5–55 disulfide directly from R and indirectly from I are approximately the same (0.028 s^{-1} vs 0.03 s^{-1}; Table IV).[4] On the basis of their electrophoretic mobilities, the one-disulfide bond intermediates comprising I are largely unfolded. Thus, the indirect mechanism does not seem to be favored significantly by conformational differences in the reduced protein or the one-disulfide bond intermediates comprising I. The *apparently* more rapid appearance of Ala30/Ala51, compared to that of (Cam14/Cam38; Ala30/Ala51), on folding in the presence of DTT$_S^S$, cannot be due to intramolecular transitions per se. Instead, because a thiol–disulfide exchange reagent is needed to mediate formation of the 5–55 disulfide, the probability of productive bond formation is increased if the reagent is intramolecular in nature; i.e., if a protein disulfide is already present. The presence of the one-disulfide bond species, labeled I, effectively increases the chance of reaction by 320 (0.03 s^{-1}/0.000093 s^{-1} M^{-1}; Table IV). There is precedent for acceleration of an apparent rate of a reaction simply by covalent incorporation of a reactant into a protein. Using an inactive mutant of aspartate aminotransferase, where the active-site general base, Lys258, was replaced with alanine, Toney and Kirsch (1989) showed that addition of exogenous primary amines could partially substitute for Lys258 in the transamidation of L-cysteinesulfinate. It was estimated that, because of its intramolecular nature, Lys258 has a rate effect 250 times that of ethylamine.

Scheme V



Relationship of the Ala30/Ala51 Folding and Unfolding Mechanisms to That of (14–38, 5–55) in the BPTI Pathway. Creighton and Goldenberg (1984) modeled the (14–38, 5–55) folding and unfolding mechanisms as two steps, both dependent on exogenous redox reagents. In the first step one-disulfide bond intermediates are formed from reduced protein, and in the second (14–38, 5–55) is formed. This formalism is consistent with the more complicated mechanism proposed for Ala30/Ala51, although values for rate constants of (5–55) formation and reduction cannot both be identical in the BPTI and Ala30/Ala51 pathways. This point is illustrated by the thermodynamic cycle shown in Scheme V which incorporates previous estimates of the DTT-dependent rate constants for formation and reduction of the one-disulfide bond intermediates (I), two-disulfide bond intermediates (II), and (14–38, 5–55), the first-order rate constants for interconversion of II, and (14–38, 5–55) (Creighton & Goldenberg, 1984) and is expanded to include interconversion between I and (5–55).[5]

[4] If all one-disulfide bond intermediates, excluding (5–55), rapidly equilibrate on a time scale of seconds and all therefore participate, directly or indirectly, in thiol–disulfide rearrangements leading to (5–55), then it can be shown that the observed rate constant (k_3) for formation of (5–55) is related to microscopic intramolecular rate constants (k_i, k_j) for forming (5–55) and equilibrium constants (K_{ij}) between the one-disulfide bond intermediate, I_i, and each of the other intermediates, excluding (5–55) as

$$k_3 = \frac{k_i + \sum k_j K_{ij}}{1 + \sum K_{ij}}$$

[5] For the reaction (5–55) ↔ (14–38, 5–55), the DTT$_S^S$-dependent forward rate constant was calculated using eq 3 and the intramolecular rate constant calculated from the third-order rate constant for reduction of (14–38, 5–55) (Creighton & Goldenberg, 1984).

For a closed cycle, the equilibrium constant for I and (5–55) must equal 0.072. This value is 4.7-fold smaller than that calculated using k_3 and k_{-3} of this work. Simulations incorporating the step I ↔ (5–55) into the BPTI folding pathway, with k_3 and a rate constant 4.7-fold larger than k_{-3} or with k_{-3} and a rate constant 4.7-fold smaller than k_3, and other rate constants determined by Creighton and Goldenberg (1984) reproduce the experimental BPTI folding and (14–38, 5–55) unfolding data. (Simulations not shown.) While the simulated rate constants for I ↔ (5–55) are not unique, the simulations are consistent with indirect formation of (5–55) from other one-disulfide bond intermediates which behave as a kinetically homogeneous population.

Finally, the rate of 14–38 disulfide breakage is about 4-fold less and the rate of its formation about 17-fold less for (14–38, 5–55) than for Ala30/Ala51. At least at this step in the mechanism, alanines at positions 30 and 51 behave more like a disulfide than like thiols. (See below.)

Comparison of the BPTI Folding Mechanism with Those of Mutants Lacking One or Two Disulfide Bonds. The redox conditions, shown in Figure 8, were used by Creighton and Goldenberg (1984) to describe the free energy profile for BPTI folding. BPTI was estimated to be 5.2 kcal/mol more stable than reduced protein for these redox conditions. We can compare relative stabilities of the various intermediates and transition states of the different folding mechanisms by assuming that reduced protein always has an apparent free energy of zero and then calculating the difference in the changes in free energies ($\Delta\Delta G$) at each step in the respective mechanisms. Values of $\Delta\Delta G$ are given in parentheses in Figure 8. If the relative free energies of the reduced proteins are not identical, then calculated $\Delta\Delta G$'s differ by some unknown constant. As we are interested in trends and relative magnitudes of the changes in ΔG, not in absolute values of the free energies, that caveat is unimportant. Only one mutant has been studied; we cannot discriminate between energetic effects due to substitution of the disulfide by alanines and loss of a cross-link.

As might be expected (Creighton & Goldenberg, 1984; Hurle et al., 1990), the mutations at positions 30 and 51 destabilize BPTI relative to reduced BPTI. Destabilizing effects are fully accounted for immediately after the rate-limiting step in folding. The (5–55) intermediate is destabilized by the same amount relative to (30–51, 5–55) in the BPTI folding pathway (or to Ser14/Ser38) as Ala30/Ala51 is to BPTI (Figure 8). Destabilization is reflected, in part, as a large increase in the rate of unfolding (at the rate-limiting step) of Ala30/Ala51 compared to BPTI (Table IV). Interestingly, the second-order rate constant of reduction of the 5–55 disulfide by DTT^{SH}_{SH} is increased in (Cam14/Cam38; Ala30/Ala51) compared to that of Ser14/Ser38 by a factor equivalent to that for the intramolecular rearrangement steps on the Ala30/Ala51 and BPTI unfolding pathways. This finding suggests that the mechanisms of direct reduction are similar in Ser14/Ser38 and (Cam14/Cam38; Ala30/Ala51) as are the mechanisms of intramolecular reduction in BPTI and Ala30/Ala51. Furthermore, the role of the 30–51 disulfide must be similar in both cases—the presence of the disulfide makes it more difficult to distort the native structure exposing the 5–55 disulfide to either intra- or intermolecular reducing reagents.

The mutations at positions 30 and 51 stabilize the rate-limiting transition states in the Ser14/Ser38 and BPTI folding pathways by similar amounts, 1.5–1.7 kcal/mol. For the Ser14/Ser38 and (Cam14/Cam38; Ala30/Ala51) transition states, the absolute value of $\Delta\Delta G$ directly reflects the 17.5-fold increase in the rate of formation of the 5–55 disulfide (Table IV). The 6-fold increase in the rate of formation of the 5–55 disulfide for the rearrangement pathways (Table IV) only accounts for about two-thirds of the net stabilization of the Ala30/Ala51 transition state. The small difference (~0.4 kcal/mol) between the absolute value of $\Delta\Delta G$ for the transition states in the Ala30/Ala51 and BPTI rearrangement pathways and the free energy change directly attributable to the folding rate acceleration might be due to experimental uncertainty. However, a second explanation can be proposed. The rate of reduction for the two-disulfide bond species, II, in the BPTI folding pathway is 7–10-fold faster than the rate of reduction for the BPTI one-disulfide bond species, I (Creighton & Goldenberg, 1984). The accelerated reduction rate of disulfides in II may be due to disulfide bond strain introduced in part by the presence of the 30–51 disulfide. If this is the case, the 30–51 disulfide is responsible for about 0.4 kcal of strain energy (also, see below).

Stabilization of the (Cam14/Cam38; Ala30/Ala51) and (Ala30/Ala51) rate-limiting transition states, relative to those of Ser14/Ser38 and BPTI, also contribute to the observed increases in (Cam14/Cam38; Ala30/Ala51) and Ala30/Ala51 unfolding. In the terminology of Beasty et al. (1986), replacement of the disulfide with alanines constitutes a mixed equilibrium–kinetic mutation since the free energies of the folded and transition states are altered relative to each other and to the reduced state. Both a peptide model for the one-disulfide bond intermediate, (30–51), and the intermediate itself have been shown to contain significant amounts of native-like secondary and tertiary structure (Oas & Kim, 1988; Kosen et al., 1981, 1983; States et al., 1987). It is tempting to speculate that the destabilization of the rate-limiting transition states of Ser14/Ser38 and BPTI are a consequence of disruption of native-like features in the folding intermediates immediately preceding these transition states.

The rates of forming and breaking the 14–38 disulfide in Ala30/Ala51 are virtually the same as those for making and breaking that disulfide in BPTI. While the transition state for this reaction is destabilized relative to the reduced protein, the extent of destabilization is the same as that of (5–55) and (Ala30/Ala51) relative to reduced protein, a direct consequence of the rate similarities for 14–38 disulfide formation and breakage in Ala30/Ala51 and BPTI.

The rate of initial disulfide formation during Ala30/Ala51 folding is somewhat less than that found for BPTI folding (Table IV). On the other hand, eliminating cysteines 14 and 38 decreases the rate of initial disulfide formation by a greater factor (Creighton, 1977a; Goldenberg, 1988; P. A. Kosen, unpublished results). Initial disulfide formation was suggested above to depend on separation of the cysteines along the chain. If this suggestion is generally true, then it would be expected that initial disulfide formation for Ala30/Ala51 would be slower than that for mutants in which cysteines 14 and 38 were eliminated because the 51–55 disulfide cannot be formed from reduced Ala30/Ala51. (Residue separation of cysteines 51 and 55 is 4; for all other possible one-disulfide bond species, residue separation is ≥8.) However, if specific conformational preferences or steric restrictions preclude formation of the 51–55 disulfide, then the relative rates of one-disulfide bond formation for Ala30/Ala51 and derivatives missing cysteines 14 and 38 are in reasonable agreement with calculations. Excluding the 51–55 disulfide, the weighted average separation of cysteine residues is greater in derivatives missing cysteines 14 and 38 than those missing cysteines 30 and 51. The one-

disulfide bond intermediate (51–55) has yet to be identified among the BPTI one-disulfide bond intermediates (Creighton, 1974b; Weissman & Kim, 1991).

Finally, reduction of the one-disulfide bond population is somewhat faster for Ala30/Ala51 than for BPTI or Ser14/Ser38. This increased rate of reduction accounts for much of the decreased stability noted in Figure 8 and may be due to disulfide bond strain in one or all the intermediates.

Comparison of Disulfide Bond-Coupled and Denaturant-Dependent Folding Mechanisms of BPTI Mutants Lacking One Disulfide Bond. Hurle et al. (1990) examined the Gdn-HCl-dependent kinetics of folding and unfolding of Ala30/Ala51 and Ala14/Ala38 using manual mixing methods and monitoring changes in ultraviolet absorbance. The denaturant-dependent folding processes reflect steps occurring after the correct disulfides are formed. Extrapolated rate constants for folding (excluding those involving possible proline peptide bond isomerizations) in the absence of Gdn-HCl correspond to rates that are much more rapid than the disulfide bond-coupled folding steps and consequently would not be observable for experiments reported here.

In denaturant-dependent folding studies, observed rates of Ala14/Ala38 folding and unfolding are faster than those of Ala30/Ala51. For the disulfide bond-coupled kinetic studies, mutants or chemical derivatives of BPTI missing the 14–38 disulfide fold more slowly than Ala30/Ala51 (Table IV; Creighton, 1977a; Marks et al., 1987b; Goldenberg, 1988; P. A. Kosen, unpublished observations). While the two sets of studies appear to be contradictory, inversion of the relative rates probably reflects conformational differences in the molecules before the rate-limiting steps of folding; different descriptions of folding and unfolding conformational transitions are needed. Before the rate-limiting folding step, the 30–51 disulfide is present in mutants and chemical derivatives of BPTI missing cysteines 14 and 38 and, obviously, missing in Ala30/Ala51. The 30–51 disulfide favors native-like interactions as noted above. Thus during denaturant-dependent folding, there is less need to search for the correct orientation of the residues in the immediate vicinity of residues 30 and 51 when the 30–51 disulfide is present. If, on the other hand, a disordered conformation is required to form the 5–55 disulfide during disulfide bond-coupled folding, the presence of the 30–51 disulfide, stabilizing the correctly folded conformation, will make it more difficult to attain a disordered state.

Acknowledgments

We thank Drs. John Altman, David Baker, Fred Cohen, Mark Hurle, and Kathy Perry for helpful discussion and Dr. Peter S. Kim for communicating results prior to publication. The UCSF Mass Spectrometry Facility (A. L. Burlingame, Director) is supported by grants from the National Institutes of Health (RR01614) and from the National Science Foundation (DIR 8700766).

Registry No. BPTI, 9087-70-1; DTT^{SH}_{SH}, 3483-12-3; DTT^{S}_{S}, 14193-38-5; GSH, 70-18-8; GSSG, 27025-41-8; Cys, 52-90-4; Ala, 56-41-7.

References

Aberth, W., Straub, K., & Burlingame, A. L. (1982) *Anal. Chem. 54*, 2029.

Altman, J. D., Henner, D., Nilsson, B., Anderson, S., & Kuntz, I. D. (1991) *Protein Eng. 4*, 593.

Beasty, A. M., Hurle, M. R., Manz, J. T., Stackhouse, T., Onuffer, J. J., & Matthews, C. R. (1986) *Biochemistry 25*, 2965.

Biemann, K. (1990) *Methods Enzymol. 193*, 886.

Chan, H. S., & Dill, K. A. (1990) *J. Chem. Phys. 92*, 3118.

Cottrell, J. S., & Evans, S. (1987) *Anal. Chem. 59*, 1990.

Creighton, T. E. (1974a) *J. Mol. Biol. 87*, 579.

Creighton, T. E. (1974b) *J. Mol. Biol. 87*, 603.

Creighton, T. E. (1975a) *J. Mol. Biol. 96*, 767.

Creighton, T. E. (1975b) *J. Mol. Biol. 96*, 777.

Creighton, T. E. (1977a) *J. Mol. Biol. 113*, 275.

Creighton, T. E. (1977b) *J. Mol. Biol. 113*, 295.

Creighton, T. E. (1977c) *J. Mol. Biol. 113*, 313.

Creighton, T. E. (1988) *Biophys. Chem. 31*, 155.

Creighton, T. E., & Goldenberg, D. P. (1984) *J. Mol. Biol. 179*, 497.

Eigenbrot, C., Randal, M., & Kossiakoff, A. A. (1990) *Protein Eng. 3*, 591.

Ellman, G. L. (1959) *Arch. Biochem. Biophys. 82*, 70.

Falick, A. M., Wang, G. H., & Walls, F. C. (1986a) *Anal. Chem. 58*, 1308.

Falick, A. M., Walls, F. C., & Laine, R. A. (1986b) *Anal. Biochem. 159*, 132.

Goldenberg, D. P. (1988) *Biochemistry 27*, 2481.

Goldenberg, D. P., & Creighton, T. E. (1984) *Anal. Biochem. 138*, 1.

Goldenberg, D. P., Frieden, R. W., Haack, J. A., & Morrison, T. B. (1989) *Nature 338*, 127.

Hurle, M. R., Marks, C. B., Kosen, P. A., Anderson, S., & Kuntz, I. D. (1990) *Biochemistry 29*, 4410.

Hurle, M. R., Eads, C. D., Pearlman, D. A., Seibel, G. L., Thomason, J., Kosen, P. A., Kollman, P., Anderson, S., & Kuntz, I. D. (1991) *Protein Sci. 1*, 91.

Jacobson, H., & Stockmayer, W. H. (1950) *J. Chem. Phys. 18*, 1600.

Kosen, P. A., Creighton, T. E., & Blout, E. R. (1981) *Biochemistry 20*, 5744.

Kosen, P. A., Creighton, T. E., & Blout, E. R. (1983) *Biochemistry 22*, 2433.

Krupka, R. M., Kaplan, H., & Laidler, K. J. (1966) *Trans. Faraday Soc. 62*, 2754.

Marks, C. B., Naderi, H., Kosen, P. A., Kuntz, I. D., & Anderson, S. (1987a) *Science 235*, 1370.

Marks, C. B., Naderi, H., Kosen, P. A., Kuntz, I. D., & Anderson, S. (1987b) in *UCLA Symposia on Molecular and Cellular Biology, New Series*, pp 335–340, A. R. Liss Inc., New York.

Oas, T. G., & Kim, P. S. (1988) *Nature 336*, 42.

Reisfeld, R. A., Lewis, U. J., & Williams, D. E. (1962) *Nature (London) 195*, 281.

Richardson, J. S. (1981) *Adv. Protein Chem. 34*, 167.

Schwarz, H., Hinz, H. J., Mehlich, A., Tschesche, H., & Wenzel, H. R. (1987) *Biochemistry 26*, 3544.

Shaked, Z., Szjewski, R. P., & Whitesides, G. M. (1980) *Biochemistry 19*, 4156.

Snyder, G. H. (1987) *Biochemistry 26*, 688.

Snyder, G. H., Cennerazzo, M. J., Karalis, A. J., & Field, D. (1981) *Biochemistry 20*, 6509.

Staley, J. P., & Kim, P. S. (1990) *Nature 344*, 685.

States, D. J., Dobson, C. M., Karplus, M., & Creighton, T. E. (1984) *J. Mol. Biol. 174*, 411.

States, D. J., Creighton, T. E., Dobson, C. M., & Karplus, M. (1987) *J. Mol. Biol. 195*, 731.

Szajewski, R. P., & Whitesides, G. M. (1980) *J. Am. Chem. Soc. 102*, 2011.

Toney, M. D., & Kirsch, J. F. (1989) *Science 243*, 1485.

Vincent, J.-P., Chicheportiche, R., & Lazdunski, M. (1971) *Eur. J. Biochem. 23*, 401.
Walls, F. C., Baldwin, M. A., Falick, A. M., Gibson, B. W., Gillece-Castro, B. L., Kaur, S., Maltby, D. A., Medzihradszky, K. F., Evans, S., & Burlingame, A. L. (1990) in *Biological Mass Spectrometry* (Burlingame, A. L., McCloskey, J. A., Eds.) p 197, Elsevier, Amsterdam.
Weissman, J. S., & Kim, P. S. (1991) *Science 253*, 1386.
Zahler, W. L., & Cleland, W. W. (1968) *J. Biol. Chem. 243*, 716.

Contributions of the Polar, Uncharged Amino Acids to the Stability of Staphylococcal Nuclease: Evidence for Mutational Effects on the Free Energy of the Denatured State†

Susan M. Green, Alan K. Meeker, and David Shortle*
Department of Biological Chemistry, The Johns Hopkins University School of Medicine, 725 North Wolfe Street, Baltimore, Maryland 21205

Received December 24, 1991; Revised Manuscript Received April 8, 1992

ABSTRACT: In order to quantitate the contributions of the polar, uncharged amino acids to the stability of the native state of staphylococcal nuclease, each of the 13 alanines, 9 glycines, 9 threonines, 6 prolines, 6 glutamines, 6 asparagines, and 3 serines was substituted, either with both alanine and glycine or with 1 of these 2 amino acids plus valine. For each mutant, the stability to reversible denaturation (ΔG_{H_2O}) was quantitated by determining the K_{app} for this reaction as a function of guanidine hydrochloride concentration. In addition, the parameter m_{GuHCl} (=d(ΔG)/d[GuHCl]) was calculated from the data. To identify the local structural features responsible for the relatively large and variable changes in ΔG_{H_2O} and m_{GuHCl} observed for the same type of substitution at different locations in nuclease, statistical correlations were sought between ΔG_{H_2O}, m_{GuHCl}, and a number of descriptors of the local structure. As with substitutions of the large hydrophobic amino acids [Shortle, D., Stites, W. E., & Meeker, A. K. (1990) *Biochemistry 29*, 8033–8041], mutation of polar, uncharged residues to Gly leads to a change in stability that, on average, correlates well with the degree to which the wild-type residue is buried. This correlation is especially significant for threonine, an amino acid with both polar and hydrophobic character, but is not demonstrated for the more typically hydrophobic residue alanine. As reported in the previous study of alanine/glycine substitutions of hydrophobic residues, a significant correlation between changes in stability and changes in the value of m_{GuHCl} is again observed, strengthening the conclusion that the putative structural changes in the denatured state which lead to increases or decreases in m_{GuHCl} are responsible for a significant fraction of the stability loss for an average mutant. The existence of this correlation is consistent with the denatured state of wild-type staphylococcal nuclease having evolved to a relatively high free energy via optimization of a balance between a maximal exposure of hydrophobic surface and a minimal gain in chain entropy. On average, mutations are less stable in proportion to the extent of which they perturb this balance. A new and puzzling correlation is reported between the extent of buriedness of a residue in the wild-type native state versus the difference in m_{GuHCl} between the Ala mutation and the Gly mutation at that position.

Analysis of mutant proteins with single amino acid substitutions provides a general strategy for quantitating the contributions of individual residues to protein folding and stability. While the actual construction and isolation of mutant proteins are often straightforward, the study of how they differ from wild-type and the quantitative interpretation of any differences are slow and uncertain processes. Perhaps the greatest obstacle to interpretation of such data is the multiplicity of mechanisms by which a mutant residue could potentially alter the free energy of the reaction. On the native side of the equation, a mutation could alter side-chain hydrophobicity, steric size and shape, electrostatic charge, α-helix and β-strand propensities, hydrogen-bonding potential, etc. On the other side of the equation, changes in chain entropy and hydrophobicity or electrostatic charge could modify the free energy of the denatured state.

One strategy for dissecting out the roles of these various interactions is to analyze several different substitutions at a site of interest and then correlate changes in stability with one or more indices that reflect the physicochemical properties of the different amino acid side chains. In this way, the most significant stabilizing interactions involving the wild-type residue at one position can sometimes be inferred (Alber et al., 1987, 1988; Yutani et al., 1987; Matsumura et al., 1988, 1989; Pakula & Sauer, 1990; Connelly et al., 1991; Stites and Shortle, unpublished data). A second, complementary strategy is to mutate each residue of a particular amino acid type one at a time and then search for those features of the local structure that best correlate with the stability change at each of the sites (Shortle et al., 1990). Whereas the first approach is designed to identify those side-chain features that contribute to stability at one particular site, the second attempts to identify those features of the local environment surrounding a class of residues which interact significantly with the wild-type side chain and thereby determine what its contribution to stability will be.

In a previous report, we described the application of this second strategy to the large hydrophobic amino acids in the small protein staphylococcal nuclease (Shortle et al., 1990).

† This work was supported by NIH Grant GM34171 (D.S.) and by NIH Postdoctoral Fellowship GM14306 (S.M.G.).

0006-2960/92/0431-5717$03.00/0