Protein Science (1993), *2*, 395–403. Cambridge University Press. Printed in the USA.

Unnatural amino acid packing mutants of *Escherichia coli* thioredoxin produced by combined mutagenesis/chemical modification techniques

R. WYNN AND F.M. RICHARDS

Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut 06511

(RECEIVED October 1, 1992; REVISED MANUSCRIPT RECEIVED November 9, 1992)

Abstract

We have produced several mutants of *Escherichia coli* thioredoxin (Trx) using a combined mutagenesis/chemical modification technique. The protein C32S, C35S, L78C Trx was produced using standard mutagenesis procedures. After unfolding the protein with guanidine hydrochloride (GdmCl), the normally buried cysteine residue was modified with a series of straight chain aliphatic thiosulfonates, which produced cysteine disulfides to methane, ethane, 1-*n*-propane, 1-*n*-butane, and 1-*n*-pentane thiols. These mutants all show native-like CD spectra and the ability to activate T7 gene 5 protein DNA polymerase activity. In addition, all mutants show normal unfolding transitions in GdmCl solutions. However, the midpoint of the transition, $[GdmCl]_{1/2}$, and the free energy of unfolding at zero denaturant concentration, $\Delta G_{(H_2O)}$, give inverse orders of stability. This effect is due to changes in m, the dependence of $\Delta G^0_{unfolding}$ on the GdmCl concentration. The method described here may be used to produce unnatural amino acids in the hydrophobic cores of proteins.

Keywords: cysteine modification; protein folding; thioredoxin; unnatural amino acids

Thioredoxin is a small monomeric protein of M_r approximately 12,000. It contains a single active-site disulfide bond, which can be reduced by NADPH and thioredoxin reductase. The reduced form of the protein can serve as a reducing agent for a variety of partners including protein disulfides, ribonucleotide reductase, and methionine sulfoxide reductase (reviewed in Holmgren, 1989). *Escherichia coli* thioredoxin is a necessary component of T7 DNA polymerase (Mark & Richardson, 1976) and is also required for the assembly of filamentous viruses f1 and M13 (Lim et al., 1985; Russel & Model, 1985). The sequences of 17 different thioredoxins are currently known (Eklund et al., 1991). Protein disulfide isomerase contains two regions with internal homology to thioredoxin (Edman et al., 1985). When the sequences for the homologous portions of these proteins are taken into account, there are 24 known primary structures. This sequence data base has allowed for the determination of which residues are highly conserved in thioredoxin (Eklund et al., 1991). In addition to the sequence information and the wealth of biochemical data regarding thioredoxin, a number of physical studies have been conducted.

The structure of the oxidized form of the *E. coli* protein has been determined by crystallographic methods (Holmgren & Soderberg, 1970; Katti et al., 1990; see Fig. 1 and Kinemage 1). It consists of two conformational domains, $\beta\alpha\beta\alpha\beta$ and $\beta\beta\alpha$. The two domains are connected by a single turn, α-helix and a 3_{10}-helix. The β-sheet forms the core of the molecule and is packed on both sides by helices. The reduced form structure has also been determined and shows little change from that determined for the oxidized form (Dyson et al., 1990).

The folding and unfolding kinetics have been examined and show multiphasic behavior, the slowest phase of which corresponds to a prolyl cis-trans isomerization (Kelley & Stellwagen, 1984; Kelley & Richards, 1987). Unfolding of thioredoxin with denaturants follows a two-state model, and the stability of several variants has been studied (Gleason et al., 1990; Langsetmo et al., 1991; Lin & Kim, 1991; Hellinga et al., 1992). In addition, tritium exchange, partial specific volume, and adiabatic compressibility measurements have been made on both the oxidized and reduced forms (Kaminsky & Richards, 1992a,b).

Reprint requests to: Dr. F.M. Richards, Department of Molecular Biophysics and Biochemistry, P.O. Box 6666, 260 Whitney Avenue, New Haven, Connecticut 06511.

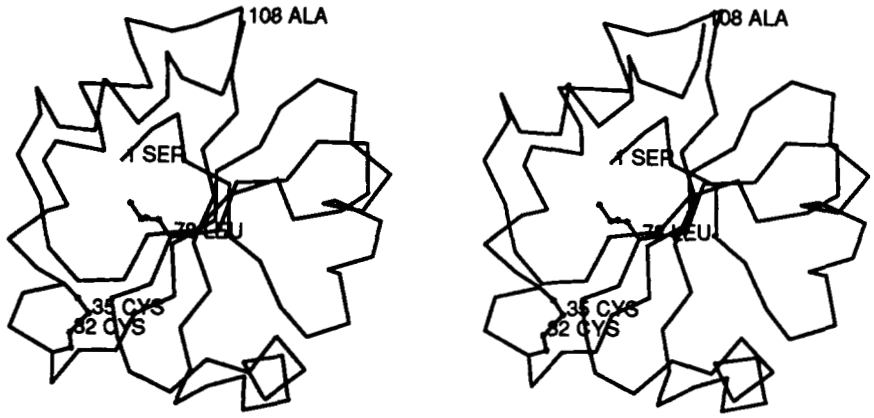


Fig. 1. Alpha-carbon trace of oxidized *Escherichia coli* thioredoxin with Leu 78 and the active-site disulfide included. Coordinates were from the structure of Katti et al. (1990). The β-sheet is shown approximately perpendicular to the viewing plane. The helices providing interior packing appear to the left and above Leu 78.

Proteins are only marginally stable at room temperature with $\Delta G_{\text{unfolding}}$ typically falling in the 5–20-kcal/mol range (Privalov, 1979). The hydrophobic effect is thought to be the main driving force for protein folding (Kauzmann, 1959; Dill, 1990). However, a number of other factors are also thought to be important, including close packing of the native structure (Richards, 1977; Ponder & Richards, 1987), electrostatics (Anderson et al., 1990; McNutt et al., 1990), hydrogen bonding (Baker & Hubbard, 1984), and intrinsic residue propensities (Hecht et al., 1986; Dao-Pin et al., 1990). One must also account for entropy loss, both local and nonlocal, upon folding (Dill, 1990). In addition, there is uncertainty with regard to the nature of the unfolded state (Dill & Shortle, 1991; Evans et al., 1991). Due to the variety of possible effects and the fact that natural amino acids usually differ in more than a single property, it is often hard to account for experimental data quantitatively.

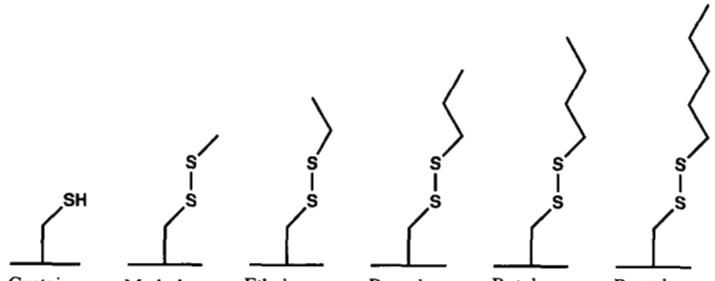
Mutagenesis is one of the most powerful techniques available for attacking the protein folding/stability problem (Matthews, 1991). While activity screens make it possible to obtain large data sets made up of sequences that produce functional and presumably stable proteins (Bowie et al., 1990), overexpression and purification of selected protein variants can provide enough material for extensive physical studies. However, mutagenesis techniques are limited to the set of 20 natural amino acids, although approaches to circumnavigate this limitation are available (Ellman et al., 1991; Mendel et al., 1992).

Prior to the development of mutagenesis techniques, chemical modification was the method of choice for producing protein variants of a chosen native protein. For many amino acids, these methods usually suffer from a lack of specificity and/or incomplete modification. In contrast, cysteine modification reagents are numerous, specific, and generally go to completion due to the intrinsic reactivity of the thiol group (Brocklehurst, 1979; Kenyon & Bruice, 1977). The relatively low abundance of cysteine in proteins (Klapper, 1977) also makes it reasonable to expect modification at one or a few sites in a protein. In this work, we have produced the *E. coli* thioredoxin (Trx) variant C32S, C35S, L78C Trx. This mutant has a single cysteine residue in a position that is normally buried in the wild-type protein. The cysteine residues in positions 32 and 35 have been changed to serines to eliminate the problem of multiple sites of chemical modification. This prevents testing the redox functions of the protein but activities not dependent on the thiol-disulfide reaction can be assayed. After unfolding in guanidine hydrochloride (GdmCl), we have modified the protein with a series of thiosulfonate reagents resulting in the formation of cysteine disulfides to methyl, ethyl, 1-*n*-propyl, 1-*n*-butyl, and 1-*n*-pentyl thiols. Upon removal of the denaturant, these modified proteins regenerate a native-like CD spectrum, are able to activate T7 gene 5 protein DNA polymerase activity, and show normal unfolding curves in GdmCl solutions. The method described here may be used for introducing many novel structures into protein hydrophobic cores.

Results

The unnatural amino acid side chains produced by the thiosulfonate modification chemistry used here are shown in Figure 2. They range from the methane thiol-like side chain of the unmodified cysteine to the methyl *n*-pentyl disulfide side chain produced using *n*-pentyl *n*-pentanethiosulfonate. The methyl disulfide side chain is closely related structurally, both in total volume and number of potential conformers, to the standard amino acid methionine. All the rest exceed any of the natural amino acids in flexibility while the pentyl derivative is substantially larger than tryptophan.

The modification chemistry was checked via three methods. First, reaction of the modified proteins with Ellman's reagent, 5,5′-dithio-*bis*-(2-nitrobenzoic acid), after unfolding in GdmCl showed no detectable thiol groups indicating complete reaction of the cysteine residue. This demonstrated that complete modification of the cysteine residue was achieved. Second, all modified pro-



No. of torsional angles	1	3	4	5	6	7
Estimated residue volume (Å^3)	118	170	196	222	248	274

Met- 3/171 Phe- 2/203 Trp- 2/238

Fig. 2. Side chains produced by thiosulfonate modification. The number of torsion angles available to alter conformations are listed along with estimates of packing volumes taken from Richards (1977). The values used were: main chain NHCHCO = 54 Å^3; H = 12 Å^3; -CH_3 = 38 Å^3; -CH_2- = 26 Å^3; -S- = 26 Å^3. For comparison, the values for methionine, phenylalanine, and tryptophan are included.

teins were analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and showed no dimer formation thus ruling out the possibility of a disulfide exchange reaction between modified protein and unmodified protein. Finally, the unfolding curves were measured following incubation with dithiothreitol (DTT). If only the desired modification reaction occurs, then incubation with DTT should regenerate the unmodified protein and thus the unfolding curve of the unmodified protein as shown in Figure 3 for the ethyl mutant. The same behavior was observed for all mutants in this study.

The CD spectra for the unmodified protein and the protein variants are presented in Figure 4. All proteins show similar spectra except for some decrease in the positive peak at approximately 198 nm for the butyl and pentyl mutants. Additionally, these spectra are comparable to CD spectra for the wild-type *E. coli* protein (Kelley & Stellwagen, 1984; Hellinga et al., 1992).

Measurement of the ability of the thioredoxin mutants to activate the DNA polymerase activity of T7 gene 5 protein was carried out at thioredoxin concentrations 300 times higher than the observed dissociation constant reported for C32S, C35S Trx and T7 gene 5 protein (Huber et al., 1986). Presumably, differences in activity are not due to changes in the binding of the mutant proteins to T7 gene 5 protein. The activity of each mutant, reported as the percent activity relative to that of reduced wild-type thioredoxin, is shown in Table 1, column 5. The unmodified protein shows the lowest activity at 32% of reduced wild-type thioredoxin. Each modified protein shows slightly higher activity ranging from 36 to 49%. The reproducibility of these assays is estimated to be ±3%.

The results from the unfolding curve measurements are also listed in Table 1. These data were compiled as described in the Materials and methods section assuming a linear relationship between the free energy of unfolding

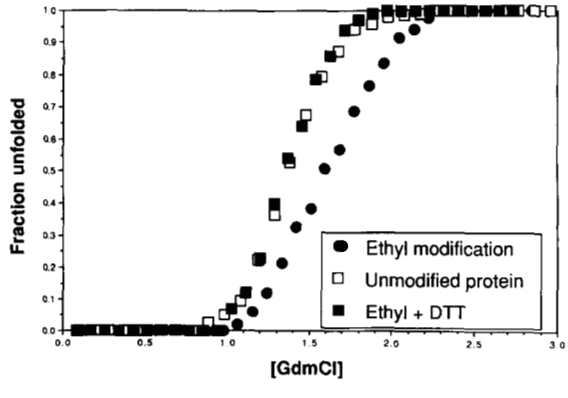


Fig. 3. Plot of the fraction of unfolded molecules as a function of GdmCl concentration. The data are for the unmodified protein corresponding to C32S, C35S, L78C thioredoxin and the ethyl modified form before and after incubation with DTT.

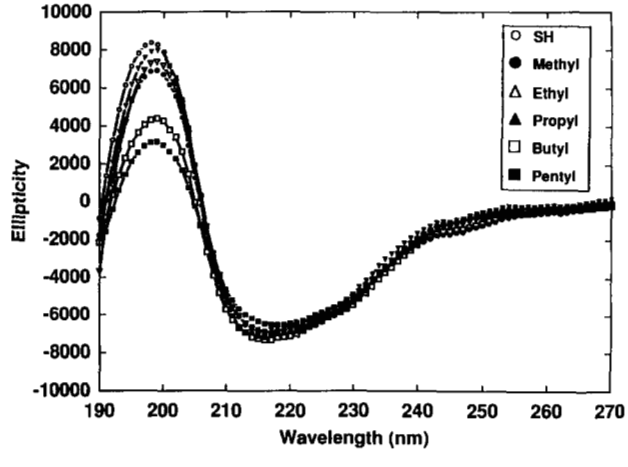


Fig. 4. CD spectra of mutant proteins.

Table 1. *Some properties of chemical variants of C32S, C35S, L78C Trx*

Mutant	$[GdmCl]_{1/2}$ [a]	m [b]	$\Delta G^0_{(H_2O)}$ [c]	Activity [d]
SH	1.37	4.25	5.84	32
Methyl	1.68	2.55	4.27	36
Ethyl	1.61	2.88	4.64	44
n-Propyl	1.54	3.44	5.30	42
n-Butyl	1.50	3.56	5.34	49
n-Pentyl	1.48	3.58	5.32	46

[a] GdmCl concentration at the midpoint of the unfolding transition expressed in molarity.

[b] Expressed in kcal mol^{-1} M^{-1}. Determined from a linear least-squares fit of $\Delta G_{unfolding}$ versus [GdmCl] as described in the text; see Equation 5.

[c] Expressed in kcal mol^{-1}; see Equation 5.

[d] Activity is the ability to activate the DNA polymerase function of phage T7 gene 5 protein. The value is expressed as the percent activity shown by reduced wild-type thioredoxin.

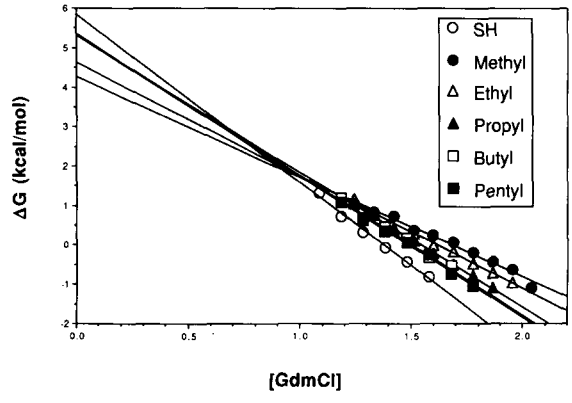


Fig. 5. $\Delta G_{unfolding}$ as a function of GdmCl concentration and extrapolation to zero denaturant concentration.

and the concentration of GdmCl. This adequately describes the data in the range of the transition as illustrated in Figure 5. Modification of the cysteine residue increases the midpoint of the transition by 0.11 M for the pentyl modification to 0.31 M for the methyl modification. All of the larger variants appear more stable than the unmodified protein by this criterion. However, the effects on $\Delta G_{(H_2O)}$ are opposite to the effects on the transition midpoint with unfolding free energy being decreased by 0.52 kcal/mol for the pentyl modification to 1.57 kcal/mol for the methyl modification. By this criterion, all of the modified proteins are less stable than the unmodified protein. As illustrated in Figure 6, there is an inverse relationship between $\Delta G_{(H_2O)}$ and $[GdmCl]_{1/2}$. This effect is due to changes in m, which goes from 4.25 kcal mol^{-1} M^{-1} for the unmodified protein to 2.55 kcal mol^{-1} M^{-1} for the methyl modified protein and then increases for the remaining mutants although it appears to be leveling off for the butyl and pentyl derivatives.

Discussion

The reagents used in this study were thiosulfonates. The chemistry of alkyl alkanethiosulfonates has been studied extensively, and the compounds have several desirable properties (Kenyon & Bruice, 1977; Brocklehurst, 1979). The reaction is independent of the nature of the alkyl groups, proceeds rapidly under mild conditions, and is selective and quantitative for thiol groups. A number of studies on a variety of proteins have borne out the effectiveness of thiosulfonate modification chemistry (Smith & Kenyon, 1974; Klineman, 1975; Nishimura et al., 1975; Smith et al., 1975; Lewis et al., 1976; Bodwell et al., 1984; Worku et al., 1984). Our results corroborate these reports. After unfolding with GdmCl, modification of the originally buried cysteine residue occurs rapidly and stoichiometrically. Furthermore, the modification is completely reversible under reducing conditions (see Fig. 3). While the thiosulfonate chemistry used here is quite sufficient, the mutagenesis/chemical modification technique described here is adaptable to any of the plentiful cysteine modification procedures.

Between 210 and 270 nm, the CD spectra of each protein are essentially identical. However, for the butyl and pentyl variants, a less positive peak is observed in the 195–200-nm region. This may indicate some local reorganization in the region of the modification. As shown in Figure 1, position 78 in the wild-type structures lies on the fourth strand of the β-sheet, is covered by two α-helices (Katti et al., 1990; Dyson et al., 1990), and is a leucine in the wild-type protein. Leucine is equivalent in size to methionine and is significantly smaller than the butyl-methyl or pentyl-methyl disulfide side chains. Protein structures frequently have a plasticity that allows for considerable adjustment to mutations (Alber, 1989). The Trx structure

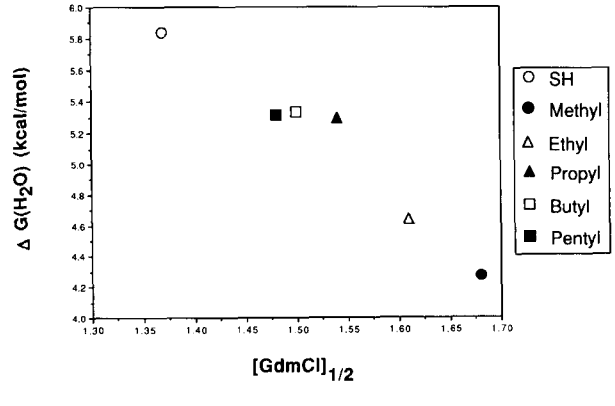


Fig. 6. $\Delta G_{(H_2O)}$ as a function of the unfolding transition midpoint, $[GdmCl]_{1/2}$.

may be able to accommodate the new side chains without significant perturbation in the secondary structure until the two largest groups are introduced (see Kinemage 2). At this point, alterations in the two helices packing on top of residue 78 may be required. Significantly decreased CD signals have also been reported for the "extreme volume" and "polar" mutants reported by Lim et al. (1992).

Previously, it has been shown that thioredoxin and T7 gene 5 protein can be reconstituted in vitro to reestablish DNA polymerase activity (Modrich & Richardson, 1975). Interaction between T7 gene 5 protein and mutant thioredoxins has also been investigated (Huber et al., 1986). This study found that C32S, C35S thioredoxin can activate gene 5 protein DNA polymerase activity to 80% of the level achieved by reduced wild-type thioredoxin but with a drastically reduced observed binding constant. However, alkylation at one or both of the active-site cysteine residues or conversion of glycine 92 to aspartate abolishes binding to gene 5 protein (Adler & Modrich, 1983). These proteins show that activation of gene 5 protein is sensitive to changes in thioredoxin. The mutants produced in this study all show the ability to stimulate gene 5 protein at levels between one-third and one-half that of reduced wild-type protein. Thus, the structure of thioredoxin in the gene 5 protein binding area is not greatly perturbed even by large changes at position 78 in agreement with the conclusion from the CD spectra.

Normally, when referring to stabilities of proteins to denaturants one reports the transition midpoint, in this case $[GdmCl]_{1/2}$, or the free energy of unfolding at zero denaturant concentration, $\Delta G_{(H_2O)}$. These two parameters usually give the same order of stability for a series of mutant proteins. This will only be true if the slope, m, relating these two parameters is the same for each protein, and typically m changes very little with amino acid substitutions. It should be noted, however, that there are numerous examples where m does change (see Dill & Shortle [1991] and references therein). In this study, an approximate inverse relationship between $\Delta G_{(H_2O)}$ and $[GdmCl]_{1/2}$ is observed (Fig. 6) as a result of large changes in m (Table 1). The change in m from the unmodified protein to the methyl cysteine disulfide mutant is among the largest reported to date for alterations at a single site.

The value of m is thought to be proportional to the difference in accessible surface area between the unfolded and native states, $m \sim (A_u - A_n)$ (Schellman, 1978), although there is not total agreement as to the mode of action for protein denaturants (Makhatadze & Privalov, 1992). Thus, changes in m can be brought about by interactions in the unfolded as well as the folded state. This reasoning has been used to explain the behavior of a number of staphylococcal nuclease mutants (Shortle & Meeker, 1986). It has also been proposed to account for the pH dependence of m for the unfolding of ribonucleases A and T1 (Pace et al., 1990).

Assuming the above relationship between m, A_u, and A_n, a decrease in m requires a decease in A_u or an increase in A_n, whereas an increase in m could be caused by either or both of the reverse changes. In changing from cysteine to cysteine methyl disulfide, the maximum side-chain area increases. Because this group is likely to be buried in the folded structure, the value of A_n should be affected little or not at all. The dramatic decrease in m from 4.25 to 2.55 would have to be reflected in A_u. Either an association of chains or a substantial contraction of a single chain would be required to produce such a large effect. It is hard to imagine how the substitution of a hydrogen atom with a CH_3S group could produce such large changes. The increase in m in going from the methyl mutant to the pentyl mutant would have to have the reverse effects. Again, the value of A_n would be expected to be unaffected or increase slightly instead of decrease and/or A_u would have to increase either through dissociation or expansion while the side chain becomes increasingly hydrophobic. Both cases seem to be counter-intuitive.

Urea gradient electrophoresis has shown that the native state of reduced thioredoxin has two conformations of different stability significantly populated (Langsetmo et al., 1989). The C32S, C35S Trx mutant was also studied and shows the same behavior. This has been attributed to cis and trans isomers at the Pro 76 peptide bond. The trans form of the oxidized protein is not significantly populated in these experiments but can be kinetically trapped. The mutants in this study should parallel the behavior of reduced thioredoxin and C32S, C35S Trx. If the cis and trans conformations of the native state have different accessible surface areas, then a change in effective surface area could be brought about by a change in the equilibrium between these two conformations. Mutations at position 78 might influence this equilibrium and thereby bring about changes in m.

It should be noted that the NMR structure of the reduced form does not show any evidence for the Pro 76 trans form of the protein. The authors suggest that low levels of urea shift the equilibrium toward the trans conformation (Langsetmo et al., 1989). If this is the case, then a linear extrapolation to zero denaturant concentration is probably not justified. Adjustment of the cis–trans equilibrium in the unfolded state could also bring about changes in surface area.

It has been shown previously that oxidized thioredoxin can incorporate 14 of the 16 amino acids tested including charged residues at position 78 (Hellinga et al., 1992). The proteins with aspartic and glutamic acid, lysine, and arginine have been overexpressed and purified. They are destabilized by 3–5 kcal/mol. It has been suggested that the charged side chains could gain access to solvent by separating the two helices that pack around this residue. A similar adjustment could take place in the case of the butyl and pentyl cysteine disulfide mutants in order to

provide room for the increased bulk of these side chains. The *m* values for the mutants with charged amino acids at position 78 have been reported to increase by up to 16% of the wild-type value (Hellinga et al., 1992). Although this is consistent with decreased clustering in the unfolded state, it may also reflect similar conformational changes in the native states of the mutant proteins in both studies.

Side-chain entropy should also play a role in the stability of these mutants as the number of conformations available to the pentyl mutant should be much higher than that for the cysteine mutant. Most or all of these conformations are likely to be frozen out in the tertiary structure. It has been proposed that side-chain entropy can account for the difference in helix-forming propensities for many amino acids in host–guest studies (Creamer & Rose, 1992). Interactions in both the native and unfolded states are affecting the unfolding behavior observed here. Further interpretation of the atypical behavior for the proteins in this study will have to await determination of the folded structures of these mutant proteins in addition to characterization of the unfolded states.

In conclusion, we have developed a technique that allows for the incorporation of unnatural amino acids into the hydrophobic cores of proteins that can be reversibly denatured. If the cysteine mutant can be expressed at high yield, the chemically derived mutants can be easily produced in high amounts. This method should prove useful for attacking a number of questions regarding protein stability.

Materials and methods

Proton NMR spectra were taken on a Bruker WM-250 spectrometer. All chemical shifts were relative to tetramethylsilane. Mass spectral analyses were carried out on an HP-5971A mass spectrometer interfaced to an HP-5990 series II gas chromatography system. Fluorometry measurements were made on a Hitachi F-3010 fluorometer. CD spectra were taken on an Aviv model 60DS CD spectropolarimeter. Methyl methanethiosulfonate was purchased from Aldrich. T7 gene 5 protein was a gift from Stanley Tabor and Young-Tae Kim.

Organic synthesis

Ethyl, *n*-propyl, and *n*-butyl disulfides were purchased from Aldrich.

n-Pentane disulfide preparation

n-Pentane disulfide was obtained from I_2 oxidation of the 1-*n*-pentane thiol. Ten milliliters of pentane thiol (0.081 mol) was dissolved in 200 mL 15% NaOH. Ten grams I_2 (0.039 mol) was added over a period of 30 min. The reaction mixture was left overnight, taken up in 200 mL ether, and washed three times with 200-mL portions of water. The ether layer was dried ($MgSO_4$) and evaporated under reduced pressure. H1 NMR ($CDCl_3$) δ 2.68(4H,T), 1.67(P,4H), 1.35(8H,m), 0.91(6H,T). Mass spec. 206(100), 136(86), 103(43), 71(84). Yield = 8.29 g (99%).

Thiosulfonate preparation

Thiosulfonates were prepared according to published procedure (Smith et al., 1975). The disulfide (0.1 mol) was mixed with 50 mL glacial acetic acid in a three-necked round-bottom flask equipped with a water-cooled condenser, dropping funnel, and thermometer. The flask was cooled to 10 °C on ice, and 21 mL of 30% H_2O_2 was added dropwise while maintaining the temperature below 15 °C. The mixture was kept at this temperature for 20 min and then slowly warmed to 60 °C. After 2 h the acetic acid was evaporated off at reduced pressure. Fifty milliliters of saturated $NaHCO_3$ was added to the residue and the product was extracted with three 30-mL portions of $CHCl_3$. The $CHCl_3$ portions were combined, dried with $MgSO_4$, and evaporated under reduced pressure. The remaining product was vacuum distilled.

Ethyl ethanethiosulfonate (EETS). H1 NMR ($CDCl_3$) δ 3.33(2H,Q), 3.16(2H,Q), 1.47(3H,T), 1.44(3H,T). Mass spec. 154(83), 95(99), 62(100). Boiling point = 71 °C (2 mm). Yield = 6.5 g (42%).

1-n-Propyl 1-n-propanethiosulfonate (PPTS). H1 NMR ($CDCl_3$) δ 3.29(2H,T), 3.12(2H,T), 1.95(2H,m), 1.78(2H,m), 1.07(6H,m). Mass spec. 182(5.7), 118(14), 76(100). Boiling point = 105 °C (2 mm). Yield = 6.0 g (33%).

1-n-Butyl 1-n-butanethiosulfonate (BBTS). H1 NMR ($CDCl_3$) δ 3.30(2H,t), 3.13(2H,t), 1.89(2H,m), 1.73 (2H,m), 1.47(4H,m), 0.96(6H,m). Mass spec. 210(3.0), 89(100), 56(79). Boiling point = 105 °C (0.15 mm). Yield = 7.32 g (32%).

1-n-Pentyl 1-n-pentanethiosulfonate (PETS). H1 NMR ($CDCl_3$) δ 3.29(2H,T), 3.11(2H,T), 1.88(2H,m), 1.72 (2H,m), 1.36(8H,m), 0.98(6H,m). Mass spec. 135(11), 103(100), 69(78), 55(15). Boiling point = 125 °C (0.1 mm). Yield = 5.85 g (25%).

Mutagenesis

Mutagenesis was carried out according to the method described by Kunkel (1985) on the vector M13mp19, which had the 0.477-kb *HincII* fragment bearing the coding region for *E. coli* thioredoxin recloned from a pUC19 construct (Kelley & Richards, 1987). It was cloned in an orientation that expressed under control of the *lac* promoter. The variant containing the cysteine codon was produced previously (Hellinga et al., 1992). A second round of mutagenesis was undertaken to replace the two native

cysteine codons with serine codons. All variants were checked by dideoxy sequencing.

Expression and protein purification

The *HincII* M13 fragment was recloned into the *HpaI* site of $pP_L\lambda$ (Pharmacia LKB technology) by blunt-end ligation using T4 DNA ligase. Ten liters of TYE media was inoculated with 400 mL of fresh culture started from a single colony. Growth was carried out at 30 °C until the absorbance at 600 nm was 0.9. The temperature was switched to 42 °C for 3.5 h. The cells were harvested and suspended in buffer A (20 mM Tris-HCl, 25 mM NaCl, pH 7.4). Following sonication, cell debris was removed by centrifugation at 12,000 × *g* for 40 min. Polyethylenimine was added to a final concentration of 0.5% (v/v). After 30 min incubation at 4 °C, the precipitate was removed by centrifugation. Protein was precipitated with 75% saturated ammonium sulfate and collected by centrifugation. After dialysis into buffer A, the sample was applied to a Bio-Gel P-30 gel filtration column and eluted in buffer A. Fractions containing thioredoxin were applied to a DE52 ion-exchange column (Pharmacia) and eluted with a 0–0.5 M NaCl gradient in buffer A. Thioredoxin was judged to be pure by SDS-PAGE. Protein concentrations were determined by absorption at 280 nm ($\epsilon_{280} = 13{,}400\ cm^{-1}$, Holmgren & Reichard [1967]). The yield of C32S, C35S, L78C Trx was 100 mg/L of culture.

Modification using thiosulfonates

Protein at a concentration of 3 mg/mL was unfolded in 3.1 M GdmCl in 50 mM potassium phosphate and incubated on ice for 1 h. In the case of modification with butyl butanethiosulfonate and pentyl pentanethiosulfonate, 30% (v/v) aqueous methanol was used instead of water to help solubilize the thiosulfonate. A twofold excess of reagent was added to the protein solution, and the mixture was incubated for 10 min. The solvent was then changed to 50 mM potassium phosphate using Centriprep 3 concentrators (Amicon). Samples were analyzed by SDS-PAGE and showed no dimer formation. To check for unreacted thiol, protein was unfolded and reacted with Ellman's reagent (5,5′-dithio-*bis*-(2-nitrobenzoic acid)) according to the following protocol. Protein was dissolved in 4.7 M GdmCl (50 mM potassium phosphate, pH 7.0) to a concentration of 0.1 mM and incubated for $\frac{1}{2}$ h. Ten microliters of Ellman's reagent stock solution (4 mg/mL) was added, and the absorbance was monitored. No detectable reaction was observed, indicating that the modification reaction proceeded to completion.

CD measurements

Measurements were made in 50 mM potassium phosphate, pH 7.0, at 25 °C. The protein concentration was 0.05 mg/mL. Measurements were taken every 1.0 nm with a 1.5-nm bandwidth using a 1-mm pathlength cuvette. A total of eight scans were averaged and a blank of buffer was subtracted.

T7 gene 5 protein DNA polymerase assays

T7 gene 5 protein was obtained from S. Tabor and Young-Tae Kim (Tabor & Richardson, 1989). T7 gene 5 protein at a concentration of 8.00×10^{-7} M and one of the thioredoxin variants at a concentration of 0.1 mM were preincubated at 0 °C in 40 mM Tris, pH 7.5, for 10 min. In the case of wild-type thioredoxin, 1 mM DTT was also included. Ten microliters of this mixture was added to 45 μL reaction buffer containing 40 mM Tris (pH 7.5), 10 mM $MgCl_2$, 0.3 mM dATP, 0.3 mM dGTP, 0.3 mM dCTP, 0.3 mM [^{3}H]dTTP (170 cpm/pmol), 0.1 mg/mL bovine serum albumin, and 0.2 mg/mL alkali-denatured calf thymus DNA (see Tabor et al. [1987] for procedure of alkali denaturization). The samples were incubated at 30 °C for 15 min and then precipitated by addition of 1.5 mL 1 M HCl–0.1 M $Na_2P_2O_7$. The DNA was collected on Whatman GF/c filter paper and washed with five 2-mL portions of 1 M HCl–0.1 M $Na_2P_2O_7$ and dried with 2 mL of ethanol. Acid-insoluble counts were determined by liquid scintillation counting. Blanks containing neither thioredoxin nor T7 gene 5 protein gave less than 0.5% of the reduced wild-type activity. Assays for T7 gene 5 protein in the absence of any thioredoxin gave approximately 3% of the reduced wild-type activity, in agreement with previous work (Tabor et al., 1987). The activity of the mutants is expressed relative to reduced wild-type thioredoxin.

Unfolding curves

Protein unfolding curves were measured in 50 mM potassium phosphate at 25 °C with a protein concentration of 10 μM. Fluorescence measurements were taken using an excitation wavelength of 295 nm. Emission was monitored at 350 nm. Guanidine hydrochloride concentrations were determined by refractive index measurements (Nozaki, 1972). The fraction of protein unfolded is given by

$$F_u = (Y_n - Y)/(Y_n - Y_u), \quad (1)$$

where Y is the fluorescence signal at a given [GdmCl] and Y_n and Y_u are the signal for the folded and unfolded proteins, respectively. Y_n and Y_u themselves are linearly dependent on [GdmCl] and were fit to the equations

$$Y_n = Y_{n,0} + a[\text{GdmCl}], \quad (2)$$

$$Y_u = Y_{u,0} + b[\text{GdmCl}], \quad (3)$$

where $Y_{n,0}$ and $Y_{u,0}$ designate the signal for the folded and unfolded protein in absence of denaturant. The unfolding equilibrium constant is related to the fraction of unfolded molecules in the following manner:

$$K_u = F_u/(1 - F_u). \quad (4)$$

Free energies are calculated using the normal logarithmic relationship ($\Delta G = -RT \ln K_u$) and the free energy of unfolding at zero denaturant is calculated from

$$\Delta G_{\text{unfolding}} = \Delta G_{H_2O} - m[\text{GdmCl}], \quad (5)$$

where only the points falling between a fraction unfolded of 0.1 and 0.9 were included. R^2 for the least-squares fit was always greater than 0.97.

After measurement of the unfolding curve for each modified protein, DTT was added to a final concentration of 1 mM. The samples were incubated for at least 12 h, and the fluorescence measurements were repeated. This serves as a check for the modification chemistry as the unfolding curve of the original unmodified protein should be regenerated.

Note added in proof

After submission of this manuscript, a paper by Lu et al. (1992) describing the application of a very similar modification procedure was brought to the attention of the authors.

Acknowledgments

We thank Dr. S. Tabor and Dr. Young-Tae Kim for the T7 gene 5 protein and Dr. H. Hellinga for helpful discussions. This work was supported by grant GM-22778 from the National Institute of General Medical Sciences.

References

Adler, S. & Modrich, P. (1983). T7 induced DNA polymerase. *J. Biol. Chem. 258*, 6956-6962.

Alber, T. (1989). Mutational effects on protein stability. *Annu. Rev. Biochem. 58*, 765-798.

Anderson, D.E., Becktel, W.J., & Dahlquist, F.W. (1990). pH induced denaturation of proteins: A single salt bridge contributes 3-5 kcal/mol to the free energy of folding of T4 lysozyme. *Biochemistry 29*, 2403-2408.

Baker, E.N. & Hubbard, R.E. (1984). Hydrogen bonding in globular proteins. *Prog. Biophys. Mol. Biol. 44*, 97-179.

Bodwell, J.E., Holbrook, N.J., & Munck, A. (1984). Sulfhydryl-modifying reagents reversibly inhibit binding of glucocorticoid-receptor complexes to DNA-cellulose. *Biochemistry 23*, 1392-1398.

Bowie, J.U., Reidhaar-Olson, J.F., Lim, W.A., & Sauer, R.T. (1990). Deciphering the message in protein sequences: Tolerance to amino acid substitutions. *Science 247*, 1306-1310.

Brocklehurst, K. (1979). Specific covalent modification of thiols: Applications in the study of enzymes and other biomolecules. *Int. J. Biochem. 10*, 259-274.

Creamer, T.P. & Rose, G.D. (1992). Side-chain entropy opposes α-helix formation but rationalizes experimentally determined helix-forming propensities. *Proc. Natl. Acad. Sci. USA 89*, 5937-5941.

Dao-Pin, S., Baase, W.A., & Matthews, B.W. (1990). A mutant T4 lysozyme (Val 131 to Ala) designed to increase thermostability by the reduction of strain with an α-helix. *Proteins Struct. Funct. Genet. 7*, 198-204.

Dill, K.A. (1990). Dominant forces in protein folding. *Biochemistry 29*, 7133-7155.

Dill, K.A. & Shortle, D. (1991). Denatured states of proteins. *Annu. Rev. Biochem. 60*, 795-825.

Dyson, H.J., Gippert, G.P., Case, D.A., Holmgren, A., & Wright, P.E. (1990). Three-dimensional solution structure of the reduced form of *Escherichia coli* thioredoxin determined by nuclear magnetic resonance spectroscopy. *Biochemistry 29*, 4129-4136.

Edman, J.C., Ellis, L., Blacher, R.W., Roth, R.A., & Rutter, W.J. (1985). Sequence of a protein disulfide isomerase and implications of its relationship to thioredoxin. *Nature 317*, 267-270.

Eklund, H., Gleason, F.K., & Holmgren, A. (1991). Structural and functional relations among thioredoxins of different species. *Proteins Struct. Funct. Genet. 11*, 13-28.

Ellman, J., Mendel, D., Anthony-CaHill, S., Noren, C., & Schultz, P.G. (1991). Biosynthetic method for introducing unnatural amino acids site specifically into proteins. *Methods Enzymol. 202*, 301-337.

Evans, P.A., Topping, K.D., Woolfson, D.N., & Dobson, C.M. (1991). Hydrophobic clustering in nonnative states of a protein: Interpretation of chemical shifts in NMR spectra of denatured states of lysozyme. *Proteins Struct. Funct. Genet. 9*, 248-266.

Gleason, F.K., Lim, C., Gerami-Nejad, M., & Fuchs, J.A. (1990). Characterization of *Escherichia coli* thioredoxins with altered active site residues. *Biochemistry 29*, 3701-3709.

Hecht, M.H., Sturtevant, J.M., & Sauer, R.T. (1986). Stabilization of λ repressor against thermal denaturation by site-directed Gly to Ala changes in α-helix 3. *Proteins Struct. Funct. Genet. 1*, 43-46.

Hellinga, H.W., Wynn, R., & Richards, F.M. (1992). The hydrophobic core of *E. coli* thioredoxin shows a high tolerance to non-conservative single amino acid substitutions. *Biochemistry 31*, 11203-11209.

Holmgren, A. (1989). Thioredoxin and glutaredoxin systems. *J. Biol. Chem. 264*, 13963-13966.

Holmgren, A. & Reichard, P. (1967). Thioredoxin 2: Cleavage with cyanogen bromide. *Eur. J. Biochem. 2*, 187-196.

Holmgren, A. & Soderberg, B. (1970). Crystallization and preliminary crystallographic data for thioredoxin from *Escherichia coli*. *J. Mol. Biol. 54*, 387-390.

Huber, H.E., Russel, M., Model, P., & Richardson, C.C. (1986). Interaction of mutant thioredoxins of *Escherichia coli* with the gene 5 protein of phage 7. *J. Biol. Chem. 261*, 15006-15012.

Kaminsky, S.M. & Richards, F.M. (1992a). Differences in hydrogen exchange behavior between the oxidized and reduced forms of *Escherichia coli* thioredoxin. *Protein Sci. 1*, 10-21.

Kaminsky, S.M. & Richards, F.M. (1992b). Reduction of thioredoxin significantly decreases its partial specific volume and adiabatic compressibility. *Protein Sci. 1*, 22-30.

Katti, S.K., LeMaster, D.M., & Eklund, H. (1990). Crystal structure of thioredoxin from *Escherichia coli* at 1.68 Å resolution. *J. Mol. Biol. 212*, 167-184.

Kauzmann, W. (1959). Some factors in the interpretation of protein denaturation. *Adv. Protein Chem. 14*, 1-63.

Kelley, R.F. & Richards, F.M. (1987). Replacement of proline-76 with alanine eliminates the slowest kinetic phase in thioredoxin folding. *Biochemistry 26*, 6765-6774.

Kelley, R.F. & Stellwagen, E. (1984). Conformational transitions of thioredoxin in guanidine hydrochloride. *Biochemistry 23*, 5095-5102.

Kenyon, G.K. & Bruice, T.W. (1977). Novel sulfhydryl reagents. *Methods Enzymol. 46*, 407-430.

Klapper, M.H. (1977). The independent distribution of amino acid near neighbor pairs into polypeptides. *Biochem. Biophys. Res. Commun. 78*, 1018-1024.

Klineman, J.P. (1975). Alkylation of two active site cysteines of yeast alcohol dehydrogenase by styrene oxide and methyl methanethiosulfonate. *Fed. Proc. Fed. Am. Soc. Exp. Biol. 34*, 600-610.

Kunkel, T.A. (1985). Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc. Natl. Acad. Sci. USA 82*, 488-492.

Langsetmo, K., Fuchs, J.A., & Woodward, C. (1989). *Escherichia coli* thioredoxin folds into two compact forms of different stability to urea denaturation. *Biochemistry 28*, 3211-3220.

Langsetmo, K., Fuchs, J.A., & Woodward, C. (1991). The conserved,

buried aspartic acid in oxidized *Escherichia coli* thioredoxin has a pKa of 7.5. Its titration produces a related shift in global stability. *Biochemistry 30*, 7603–7609.

Lewis, S.D., Johnson, F.A., & Shafer, J.A. (1976). Potentiometric determination of the ionizations at the active site of papain. *Biochemistry 15*, 5009–5017.

Lim, C.J., Haller, B., & Fuchs, J.A. (1985). Thioredoxin is the bacterial protein encoded by *fip* that is required for filamentous bacteriophage f1 assembly. *J. Bacteriol. 161*, 799–802.

Lim, W.A., Farruggio, D.C., & Sauer, R.T. (1992). Structural and energetic consequences of disruptive mutations in a protein core. *Biochemistry 31*, 4324–4333.

Lin, T. & Kim, P.S. (1991). Evaluating the effects of a single amino acid substitution on both the native and denatured states of a protein. *Proc. Natl. Acad. Sci. USA 88*, 10573–10577.

Lu, J., Baase, W.A., Muchmore, D.C., & Dahlquist, F.W. (1992). Protein folding: Assignment of the energetic changes of reversible chemical modifications to the folded or unfolded states. *Biochemistry 31*, 7765–7772.

Makhatadze, G.I. & Privalov, P.L. (1992). Protein interactions with urea and guanidinium chloride. *J. Mol. Biol. 226*, 491–505.

Mark, D.F. & Richardson, C.C. (1976). *Escherichia coli* thioredoxin: A subunit of bacteriophage T7 DNA polymerase. *Proc. Natl. Acad. Sci. USA 73*, 780–784.

Matthews, B.W. (1991). Mutational analysis of protein stability. *Curr. Opin. Struct. Biol. 1*, 17–21.

McNutt, M., Mullins, L.S., Raushel, F.M., & Pace, C.N. (1990). Contribution of histidine residues to the conformational stability of ribonuclease T1 and mutant glu58 to ala. *Biochemistry 29*, 7572–7576.

Mendel, D., Ellman, J.A., Chang, Z., Veenstra, D.L., Kollman, P.A., & Schultz, P.G. (1992). Probing protein stability with unnatural amino acids. *Science 256*, 1798–1802.

Modrich, P. & Richardson, C.C. (1975). Bacteriophage T7 deoxyribonucleic acid replication in vitro. *J. Biol. Chem. 250*, 5508–5514.

Nishimura, J.S., Kenyon, G.L., & Smith, D.J. (1975). Reversible modification of the sulfhydryl groups of *E. coli* succinic thiokinase with methanethiolating reagents. *Arch. Biochem. Biophys. 170*, 461–467.

Nozaki, Y. (1972). The preparation of guanidine hydrochloride. *Methods Enzymol. 26*, 43–51.

Pace, C.N., Laurents, D.V., & Thomson, J.A. (1990). pH dependence of the urea and guanidine hydrochloride denaturation of ribonuclease A and ribonuclease T1. *Biochemistry 29*, 2564–2572.

Ponder, J.W. & Richards, F.M. (1987). Tertiary templates for proteins: Use of packing criteria in the enumeration of allowed sequences for different structural classes. *J. Mol. Biol. 193*, 775–791.

Privalov, P.L. (1979). Stability of proteins. *Adv. Protein Chem. 33*, 167–241.

Richards, F.M. (1977). Areas, volumes, packing, and protein structure. *Annu. Rev. Biophys. Bioeng. 6*, 151–176.

Russel, M. & Model, P. (1985). Thioredoxin is required for filamentous phage assembly. *Proc. Natl. Acad. Sci. USA 82*, 29–33.

Schellman, J.A. (1978). Solvent denaturation. *Biopolymers 17*, 1305–1322.

Shortle, D. & Meeker, A.K. (1986). Mutant forms of staphylococcal nuclease with altered patterns of guanidine hydrochloride and urea denaturation. *Proteins Struct. Funct. Genet. 1*, 81–89.

Smith, D.J. & Kenyon, G.L. (1974). Non-essentiality of the active sulfhydryl group of rabbit muscle creatine kinase. *J. Biol. Chem. 249*, 3317–3318.

Smith, D.J., Maggio, E.T., & Kenyon, G.L. (1975). Simple alkanethiol groups for temporary blocking of sulfhydryl groups in enzymes. *Biochemistry 14*, 766–771.

Tabor, S., Huber, H.E., & Richardson, C.C. (1987). *Escherichia coli* thioredoxin confers processivity on the DNA polymerase activity of the gene 5 protein of bacteriophage T7. *J. Biol. Chem. 262*, 16212–16223.

Tabor, S. & Richardson, C.C. (1989). Selective inactivation of the exonuclease activity of bacteriophage T7 DNA polymerase by *in vitro* mutagenesis. *J. Biol. Chem. 264*, 6447–6458.

Worku, Y., Luzio, J.P., & Newby, A.C. (1984). Identification of histidyl and cysteinyl residues essential for catalysis by 5′-nucleotidase. *FEBS Lett. 167*, 235–240.