# The Effects of Hydrophilic to Hydrophobic Surface Mutations on the Denatured State of Iso-1-cytochrome *c*: Investigation of Aliphatic Residues†

Lynn Herrmann,‡ Bruce E. Bowler,*,‡ Aichun Dong,§ and Winslow S. Caughey§

*Department of Chemistry, University of Denver, Denver, Colorado 80208, and Department of Biochemistry and Molecular Biology, Colorado State University, Fort Collins, Colorado 80523*

*Received August 25, 1994; Revised Manuscript Received November 21, 1994*⊗

ABSTRACT: A series of hydrophilic to hydrophobic surface mutations were prepared at the highly solvent-exposed lysine 73 of iso-1-cytochrome *c* to assess the ability of such mutants to affect the energetics of the denatured state. In this report, the aliphatic hydrophobics (leucine, isoleucine, valine, alanine, glycine) were studied. The thermodynamic stability of each of these mutants was determined by guanidine hydrochloride denaturation. Both the free energy of unfolding in the absence of denaturant, $\Delta G^{\circ}_{u}{}^{H_2O}$, and the slope, *m*, of a plot of the free energy of unfolding, $\Delta G^{\circ}_{u}$, versus [guanidine hydrochloride] show significant negative correlations with the 1-octanol to water transfer free energy, $\Delta G_{tr}$, of the amino acid side chain at position 73. A negative correlation with hydrophobicity is consistent with these mutants leading to more extensive hydrophobic clustering in the denatured state, consistent with the predictions of heteropolymer theory for compact denatured states; an effect operating on the native state energetics should produce a positive correlation of $\Delta G^{\circ}_{u}{}^{H_2O}$ with hydrophobicity. Infrared amide I spectroscopy indicated native state structural perturbations for the glycine 73 and isoleucine 73 mutants. A moderate correlation of $\Delta G^{\circ}_{u}{}^{H_2O}$ was also found with α-helix propensity, suggesting that both hydrophobic effects acting on the denatured state and α-helix propensity are affecting the $\Delta G^{\circ}_{u}{}^{H_2O}$ values for these mutants.

Recent studies on protein stability suggest that while a random coil may be a useful description of the denatured state for some proteins, it is by no means a general description of the denatured state of a protein (Shortle, 1993; Dill & Shortle, 1991). Single amino acid substitutions have been found to modify interactions in the denatured state with the protein staphylococcal nuclease (Shortle & Meeker, 1986). An engineered disulfide in iso-1-cytochrome *c* has also been shown to produce a more compact denatured state, leading to stabilization of this state (Betz & Pielak, 1992). Spectroscopic studies have increasingly supported the idea that the denatured state is not a random coil (Neri et al., 1992; Sosnik & Trewhella, 1992; Seshadri et al., 1994; James et al., 1992; Flanagan et al., 1992; Alexandrescu et al., 1994). In fact, a study on the α-subunit of tryptophan synthetase demonstrates that fluorescence and circular dichroism (CD)[1] techniques which have traditionally been used to demonstrate total loss of ordered structure are not adequate to detect all forms of residual structure (Saab-Rincon et al., 1993). Recent studies involving NMR (Shortle & Abeygunawardana, 1993) and infrared spectroscopy (Bowler et al., 1994) have demonstrated mutationally-induced shifts in the conformational populations of the denatured state that were previously predicted from thermodynamic data.

In our laboratory, we have been attempting to develop rational methods to generate mutations which selectively affect the denatured state of a protein. In one approach, highly solvent-exposed amino acids are chosen for mutation, reasoning that such mutations should minimally affect the energetics of the native state of the protein. Surface-exposed amino acids generally remain mobile and interact weakly with the remainder of the native state protein structure (Shortle, 1989). Electrostatic interactions are conceivably a concern; however, work on T4 lysozyme has indicated a surprising lack of sensitivity of protein stability to dramatic changes in the surface charge properties of a protein and has demonstrated the difficulty in engineering specific charge interactions on the surface of a protein (Dao-pin et al., 1991a,b). The solvent-exposed surface sites are then mutated, converting hydrophilic residues into hydrophobic amino acids, so as to enhance hydrophobic cluster formation in the denatured state. The effects of such a mutational strategy on the free energy of unfolding, $\Delta G^{\circ}_{u}$, have been predicted by heteropolymer theory for proteins with compact denatured states (Shortle et al., 1992).

Our initial work involved replacement of the highly solvent-exposed lysine 73 of iso-1-cytochrome *c* with large aromatic or polarizable residues (Bowler et al., 1993). The mutations (Met, Tyr, Phe, and Trp) appeared to stabilize the denatured state through increased hydrophobic clustering within a more compact denatured state ($\Delta G^{\circ}_{u}{}^{H_2O}$ decreased and the slope, *m*, of $\Delta G^{\circ}_{u}$ vs [gdnHCl] decreased in magnitude). A similar inverse correlation between hydrophobicity and $\Delta G^{\circ}_{u}{}^{H_2O}$, termed a reverse hydrophobic effect, has been observed previously for replacements at a surface-exposed site in the λ Cro protein (Pakula & Sauer, 1990). To further explore the effect of mutation at position 73, this study investigates the effects of mutation of lysine 73 to aliphatic residues (Gly, Ala, Val, Leu, and Ile) on the

† This work was supported in part by NSF Grant MCB-9304751 (B.E.B.), by Colorado Agricultural Experimental Station Project 643 (W.S.C.), and by a gift from Strohtech, Inc. (W.S.C.).

* Address correspondence to this author.

‡ University of Denver.

§ Colorado State University.

⊗ Abstract published in *Advance ACS Abstracts,* February 1, 1995.

[1] Abbreviations: gdnHCl, guanidine hydrochloride; FTIR, Fourier transform infrared; CD, circular dichroism; $\Delta G^{\circ}_{u}$, free energy of unfolding; $\Delta G^{\circ}_{u}{}^{H_2O}$, free energy of unfolding in the absence of denaturant; $A_d$, solvent-exposed surface area of the denatured state; $A_n$, solvent-exposed surface area of the native state; $ddH_2O$, deionized distilled water.

0006-2960/95/0434-3040$09.00/0 

unfolding free energy of iso-1-cytochrome *c*. One issue of interest in the current work is whether the polarizable aromatic versus the nonpolarizable aliphatic amino acids differ in their ability to induce hydrophobic clusters in the denatured state. This work is therefore a further test of the generality of heteropolymer theory (Dill & Shortle, 1991; Shortle et al, 1992).

Mutations at surface-exposed amino acids have been used to study α-helix propensity in a protein context (Blaber et al., 1993, 1994; Serrano et al., 1992; Horovitz et al., 1992; Pinker et al., 1993). Two of the mutants in this study show native state structural perturbations by FTIR spectroscopy. Since position 73 is in a short stretch of α-helix, the possibility of α-helix propensity effects will be discussed. On the whole, the data presented here are consistent with an important effect of these mutations on denatured state energetics, in accord with the predictions of heteropolymer theory. However, there are indications that native state effects may be modulating the overall observed free energy of unfolding.

## EXPERIMENTAL PROCEDURES

Oligonucleotides were obtained from Operon Technologies or were a gift from Somatogen, Inc. (Boulder, CO). Site-directed mutagenesis was performed using the *in vitro* mutagenesis kit version 2.1 from Amersham Corp. Sequencing reactions were carried out using Sequenase version 2.0 from U.S. Biochemical. Helper phage R408 was obtained from Promega. Restriction enzymes were from New England Biolabs. For DNA isolations, columns and buffers from Qiagen, Inc., were utilized. Materials for growth media were from Difco. YPD, SD Leu$^-$, and YPG growth media were used for the growth of yeast (Sherman et al., 1986). All solutions were made with glass-distilled water from a Corning MP-6A Mega-Pure System running off a deionized water source. Chemicals used for buffers were all reagent grade or better.

*Preparation and Characterization of Iso-1-cytochrome c Variants.* All site-directed mutations to the iso-1-cytochrome gene (CYC1) were prepared and characterized as previously described (Bowler et al., 1993) in the phagemid pRS425/CYC1 built from the yeast shuttle vector pRS425 (Christianson et al., 1992) and the CYC1 gene (Montgomery et al., 1978). Mutagenic oligonucleotides had the general sequence: 5′-CCAGGAATATA,XYZ,CTTTGGGTTAGTC-3′ where X, Y, and Z stand for the different bases used at each site in the position 73 codon. For glycine, alanine, and valine, the oligonucleotide mixture with X;Y;Z equal to A;G,C,A;C was used. For leucine, isoleucine, and arginine, X;Y;Z were, respectively, C;A;A, G;A;T, and T;C;T. These codons represent the preferred codon usage of *Saccharomyces cerevisiae* (Ikemura, 1982; Bennetzen & Hall, 1982). Iso-1-cytochrome *c* variants were isolated, according to previous methods (Bowler et al., 1993), from GM-3C-2 *S. cerevisiae* cells [cytochrome *c* deficient; see Faye et al. (1981)] transformed with the appropriate pRS425/CYC1 plasmid. Fourier transform infrared spectra were acquired and analyzed as previously described (Bowler et al., 1993; Dong et al., 1990, 1992a). Buffer conditions for infrared spectroscopy were 50 mM sodium phosphate, pH 7.2, 1 mM EDTA, and protein concentrations were 12−20 mg/mL for measurements in a 6 μM path-length cell.

*Guanidine Hydrochloride Denaturations.* Ferricytochrome *c* was used in all experiments. Oxidized protein was separated from the oxidant $K_3Fe(CN)_6$ by G-25 chromatography using 20 mM Tris, pH 7.5, 40 mM NaCl as the running buffer. Unfolding as a function of guanidine hydrochloride concentration, [gdnHCl], was monitored by CD spectroscopy at 220 nm. Measurements were made with a JASCO 500C spectropolarimeter using a time constant of 2 s, a sensitivity of 2 mdeg/cm, and a slit width of 1800 μm using protein concentrations in the range of 1.5−2.1 μM. The guanidine hydrochloride stock was 6.0 M; its concentration was evaluated by refractive index measurements versus deionized distilled water at 25 °C prior to each experiment (Nozaki, 1972). All measurements were made at 25 ± 0.1 °C using a jacketed cell attached to a Neslab Model RTE 5 circulating bath. Two different methods were used to carry out the gdnHCl unfoldings. The first has been described previously (Bowler et al., 1993). The second method is a modification of a method developed by Shortle and Meeker (1986). Two stock solutions were required for this method: 6.0 M gdnHCl in 20 mM Tris-HCl, pH 7.5, 40 mM NaCl and a 2×-concentrated cytochrome *c* stock in the same buffer. The initial sample was prepared by mixing equal volumes of the 2×-concentrated cytochrome *c* stock and 20 mM Tris-HCl, pH 7.5, 40 mM NaCl. To increase the gdnHCl concentration, 2*x* microliters of the sample was removed from the sample cell and replaced with *x* microliters of 2×-concentrated cytochrome *c* and *x* microliters of 6.0 M gdnHCl. The value of *x* for each increase in gdnHCl concentration was calculated using a spreadsheet program. The sample was stirred for 1 min after the addition of the reagents to achieve thorough mixing. The sample compartment of our JASCO 500C has been modified to hold a stirrer so that the sample does not have to be moved between measurements. The second method produces high-quality pre- and posttransition base lines because the concentration of cytochrome *c* present for each measurement is less susceptible to pipeting errors. This method is thus a significant improvement over the first method. Direct comparison of the two methods demonstrates that both produce similar thermodynamic parameters for the same protein (Godbole and Bowler, unpublished data; Sarisky and Bowler, unpublished data).

*Denaturation Analysis.* Analysis of denaturation curves followed standard procedures (Pace, 1986) as described previously (Bowler et al., 1993). At least three separate trials were performed for each cytochrome *c* variant, and the data were merged and fit to the equation $\Delta G^\circ_u = \Delta G^\circ_u{}^{H_2O} - m[\text{gdnHCl}]$ using a weighted least-squares fit as previously described (Bowler et al., 1993). The standard deviations in slopes and intercepts were calculated using the equation $SD = \{\sum[(\delta F/\delta \Delta G_i)^2 SD(\Delta G_i)^2]\}^{1/2}$, where *F* is the function for the weighted least-squares slope or intercept and $SD(\Delta G_i)$ is the standard deviation of $\Delta G^\circ_u$ at each [gdnHCl] (Shoemaker et al., 1974). The midpoint of the unfolding transition was calculated as $[\text{gdnHCl}]_{1/2} = \Delta G^\circ_u{}^{H_2O}/m$. The error in $[\text{gdnHCl}]_{1/2}$, $\epsilon([\text{gdnHCl}]_{1/2})$, was calculated using an equation for propagation of systematic errors: $\epsilon([\text{gdnHCl}]_{1/2}/[\text{gdnHCl}]_{1/2} = \{\epsilon(\Delta G^\circ_u{}^{H_2O})/(\Delta G^\circ_u{}^{H_2O}) - \epsilon(m)/(m)\}$, since the signs of the errors in $\Delta G^\circ_u{}^{H_2O}$ and *m* are coupled (Shoemaker et al., 1974).

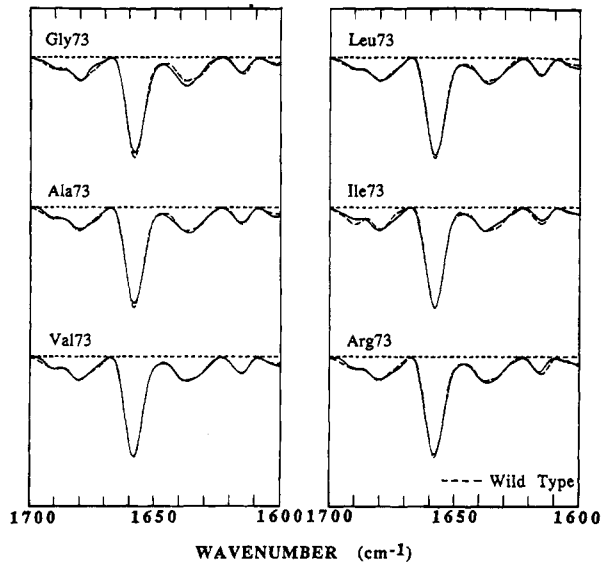


FIGURE 1: Second-derivative amide I infrared spectra for the position 73 mutant iso-1-cytochromes *c*. The spectrum of the wild-type protein (dashed line) is shown along with each mutant spectrum.

## RESULTS

To determine whether or not making mutations at the lysine 73 site disturbed the native protein structure, FTIR amide I analysis was performed. The protein amide I band (1700−1620 $cm^{-1}$) arises mainly from the C=O stretch of amide groups in the peptide bonds (Susi & Byler, 1986; Krimm & Bandekar, 1986). Each type of secondary structure, α-helix, β-sheet, etc., gives a different C=O stretch frequency as a result of differences in molecular geometry and hydrogen bonding involving the C=O and NH moieties of amide groups. Use of the second derivative of the amide I spectrum has proven to be sufficiently sensitive to permit the detection of subtle changes in secondary structure. For example, small conformational changes due to alterations in the oxidation state of iron in cytochromes *c* (Dong et al., 1992a; Bowler et al., 1993; Schlereth & Mantele, 1993), calcium binding to pentraxins (Dong et al., 1992b, 1994), and ligand binding to hemoglobins and myoglobins (Dong & Caughey, 1994; Dong et al., 1995) have been demonstrated effectively. Figure 1 shows the second-derivative amide I spectra of various mutations at position 73 (glycine, alanine, valine, leucine, isoleucine, and arginine) plotted against the spectrum of the wild-type protein. The very small differences, if any, between the second-derivative spectra of wild type and the alanine, leucine, valine, and arginine variants provide direct evidence that the differences in secondary structure among these proteins must be very small. However, the spectra of the glycine and isoleucine mutants suggest there may be significant, if small, differences in their secondary structure compared to the wild-type protein. The glycine mutant absorbance is lower at 1658 $cm^{-1}$ and greater at 1636 $cm^{-1}$ than is found for the wild-type protein, consistent with less α-helix and more β-sheet structure, respectively. The isoleucine mutant spectrum supports more turn structure (1675 $cm^{-1}$) and less β-sheet (1636 $cm^{-1}$). The above secondary structure assignments are from empirical studies reported in the literature (Susi & Byler, 1986; Dong et al., 1992a; Dong & Caughey, 1994) which are consistent with the predicted effects of transition dipole coupling on the resonance frequency of the amide I band for different types of secondary structure (Krimm & Bandekar, 1986).

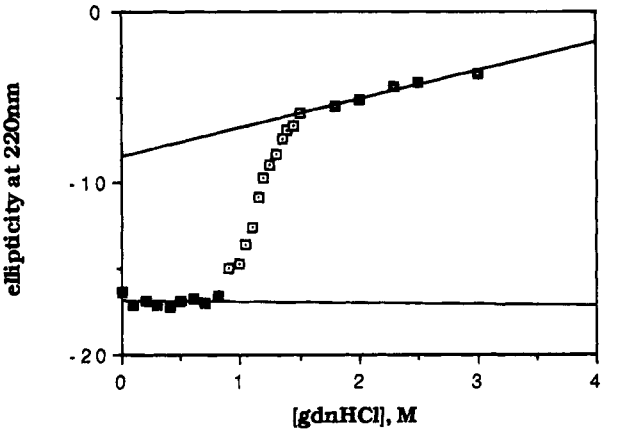


FIGURE 2: Example of a guanidine hydrochloride unfolding titration curve for leucine 73 iso-1-cytochrome *c* used to determine the free energy of unfolding in the transition region. The ellipticity in millidegrees at 220 nm is measured as a function of guanidine hydrochloride concentration. The upper and lower baselines denote the denatured and native states respectively.

Thermodynamic analysis of protein unfolding requires that a reversible process is being observed and that measurements are being made under equilibrium conditions. Previous investigators have shown that iso-1-cytochrome *c* reaches folding equilibrium within 5 min even for slow unfolding species (Ramdas & Nall, 1986; Nall, 1990; Bowler et al., 1993), well within the allotted time frame of our experiments. The question of reversibility was addressed by diluting samples of iso-1-cytochrome *c* from 3 M gdnHCl to 0.3−0.4 M gdnHCl and observing the recovery of ellipticity at 220 nm. Typically, reversibility was 94−100% (Bowler et al., 1993).

Unfolding titrations were conducted for each of the mutants, and typical sigmoidal curves were obtained by plotting [gdnHCl] versus ellipticity at 220 nm as in Figure 2. The free energy of unfolding, $\Delta G^\circ_u$, was calculated in the transition zone and plotted as $\Delta G^\circ_u$ versus [gdnHCl] (see Figure 3). The data for each mutant are plotted against data for the wild-type protein and the Trp 73 mutant [most compact denatured state from the previous study; see Bowler et al. (1993)]. The thermodynamic parameters collected in Table 1 were obtained from the plots in Figure 3, using the equation:

$$\Delta G^\circ_u = \Delta G^{\circ\,H_2O}_u - m[\text{gdnHCl}]$$

where $\Delta G^{\circ\,H_2O}_u$ is the free energy of unfolding in the absence of denaturant. The slope of this equation, *m*, is proportional to the difference in solvent-exposed surface area between the denatured and native states (Schellman, 1978):

$$\Delta A = A_d - A_n$$

The slope, *m*, can thus be taken as an indicator of the compactness of the denatured state (Dill & Shortle, 1991).

Inspection of Figure 3 and Table 1 shows that the Leu 73 mutant is most like the wild-type (Lys 73) protein, whereas the Ile 73 and Val 73 mutants are most like the compact denatured state of the Trp 73 mutant. The Ala 73 and Gly 73 mutants fall in between the wild type and Trp 73 protein

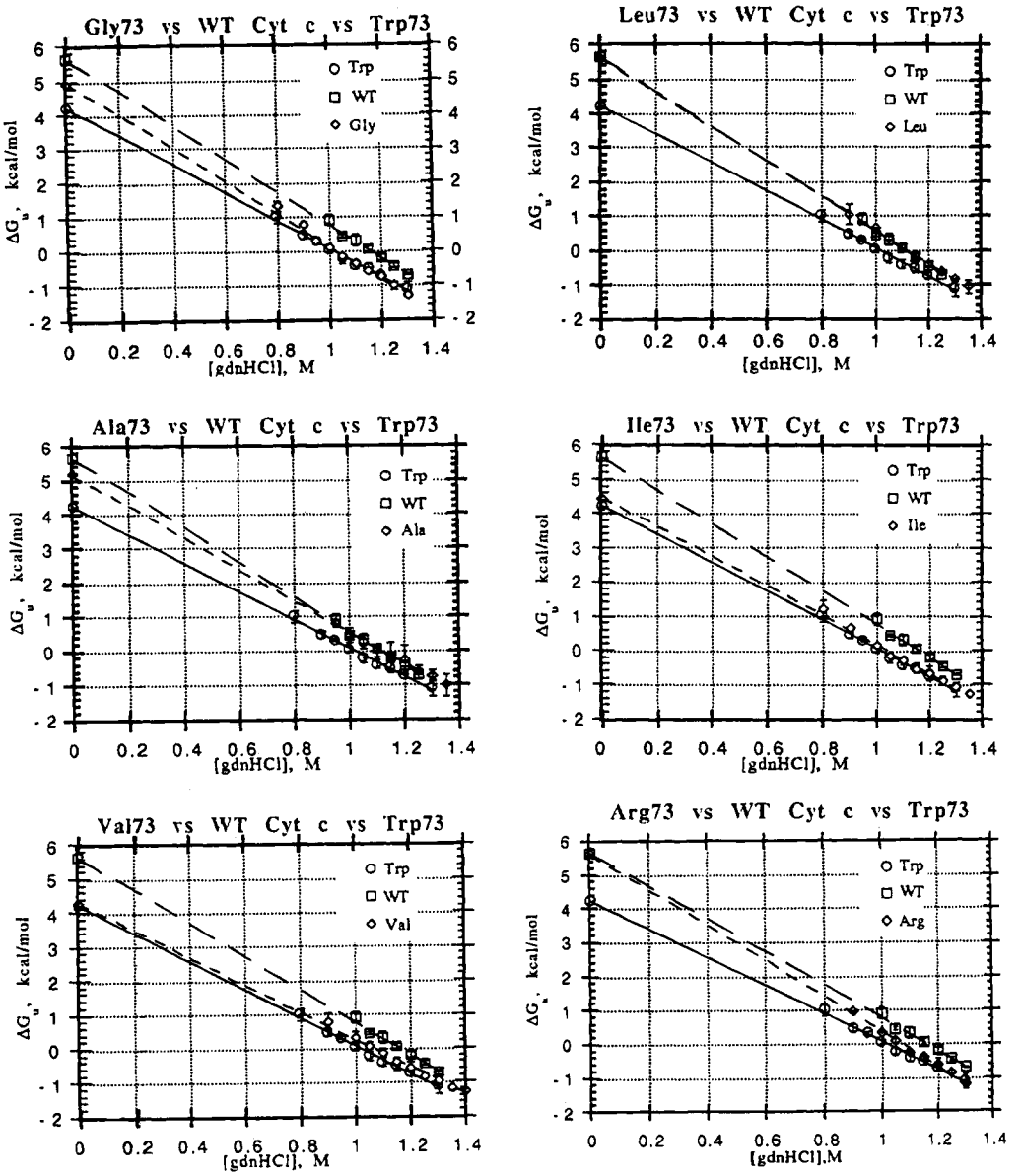


FIGURE 3: Free energy plots, $\Delta G°_u$ versus [gdnHCl], shown for Gly, Ala, Val, Leu, Ile, and Arg. Three lines which are a weighted least-squares fit are shown in each graph: wild-type (Lys 73), the mutant, and tryptophan 73. The error bars represent one standard deviation. The $y$-intercept is the $\Delta G°_u{}^{H_2O}$ value for each mutant.

Table 1: Thermodynamic Data for Iso-1-cytochrome $c$ Mutants

| cytochrome $c$ (C102S) | $\Delta G°_u{}^{H_2O}$ (kcal/mol) (25 °C) | $[gdnHCl]_{1/2}$ (M) | $m$ [kcal/(mol·M)] |
|---|---|---|---|
| wild type | 5.66 ± 0.41 | 1.16 ± 0.008 | 4.87 ± 0.32 |
| glycine 73 | 4.92 ± 0.41 | 1.07 ± 0.003 | 4.58 ± 0.37 |
| alanine 73 | 5.21 ± 0.37 | 1.17 ± 0.03 | 4.44 ± 0.43 |
| valine 73 | 4.28 ± 0.30 | 1.07 ± 0.01 | 4.01 ± 0.24 |
| leucine 73 | 5.63 ± 0.08 | 1.13 ± 0.0002 | 4.98 ± 0.07 |
| isoleucine 73 | 4.47 ± 0.12 | 1.04 ± 0.001 | 4.29 ± 0.12 |
| arginine 73 | 5.63 ± 0.11 | 1.07 ± 0.002 | 5.25 ± 0.11 |

in their properties. The Arg 73 protein has an increased slope $m$ relative to the wild-type protein, and although the $\Delta G°_u$ versus [gdnHCl] data do not overlay the wild-type data (Figure 3) as with the Leu 73 mutant data, the stability in the absence of gdnHCl, $\Delta G°_u{}^{H_2O}$, is almost identical to the wild-type protein.

## DISCUSSION

Guanidine-induced denaturation of proteins has been studied for many years to gain information about the native state folding interactions occurring within a protein. The unfolding equilibrium of many proteins can be treated as a reversible equilibrium process occurring between two states, the native and denatured (Pace, 1986):

$$N \leftrightarrow D$$

Interest in the role of the denatured state in controlling this equilibrium has increased recently, denatured states often being classed into two structural forms: compact and unfolded (Dill & Shortle, 1991). Increasingly, it has been accepted that the denatured state is not necessarily composed of random structures, but often maintains residual structural elements within the context of compact regions (Shortle, 1993; Dill & Shortle, 1991). In support of initial thermodynamic results, various investigators have provided spectroscopic proof for compact denatured states and residual secondary structure in denatured proteins using NMR spectroscopy (Neri et al., 1992; Shortle & Abeyhunawardana, 1993), fluorescence spectroscopy (James et al., 1992; Tsong, 1974), small-angle X-ray scattering (Flanagan et al., 1992),

FTIR spectroscopy of gdnHCl-denatured wild-type and mutant iso-1-cytochromes *c* (Bowler et al., 1994), and FTIR spectroscopy of thermally-denatured RNase A (Sosnick & Trewhella, 1992; Seshadri et al., 1994).

Some concern with regard to specific ion binding effects on the denaturation of proteins in the presence of gdnHCl has been raised recently (Mayr & Schmid, 1993; Hagihara et al., 1994). In the case of horse heart cytochrome *c* (Hagihara et al., 1994), progressive acylation of lysine residues increasingly destabilized the protein to urea denaturation. However, no destabilization was observed for denaturation by gdnHCl. Specific stabilization by the guanidinium ion of charge repulsion between negatively charged side chains normally shielded by lysines was proposed. Since the mutations described here involve replacement of a lysine side chain, such specific guanidinium ion effects are a potential concern. The charge isolation of the lysine 73 side chain (*vide infra*) would argue against such an effect. The variation in $\Delta G^{\circ}_{u}{}^{H_2O}$ and *m* values among the variants described here and previously (Bowler et al., 1993) argues minimally that even if a charge shielding effect is occurring, we are still observing the specific effect of each side chain not due to a putative guanidinium ion binding effect. These effects are also very species-specific, being observed for RNase T1 but not RNase A (Mayr & Schmid, 1993). We note that chloride ion has been observed to stabilize the native state of iso-1-cytochrome *c* (Hickey et al., 1988; Tsong, 1975). We observe a slight increase in α-helical structure by CD in our unfolding titrations for both wild-type and mutant proteins at low concentrations of gdnHCl (see Figure 2). The consistency of this effect for wild type and variants suggests that it can be treated as a background effect occurring due to chloride ion binding at a remote site not important to evaluating the differences between the mutants described here.

A major premise for this investigation is that by choosing a highly solvent-exposed site, the mutation at this site would not affect the native state extensively and thus perturbations to the free energy of unfolding could be attributed largely to the denatured state. Evidence for minimal native state effects is shown in the FTIR spectra in Figure 1. When lysine 73 is replaced by Ala, Leu, Val, and Arg, native state IR spectra essentially identical to the wild-type protein are observed in the structure-sensitive amide I region. Similarly, FTIR studies on the large aromatic residues, Trp, Tyr, and Phe, at this site showed negligible effects on the native state structure (Bowler et al., 1993). The Ile and Gly mutants, however, do show significant deviations from the native state structure of the wild-type protein. The probable changes in secondary structure due to these mutations have been outlined under Results.

There is still significant debate regarding the relative contributions of native and denatured state energetics to changes in unfolding free energy caused by mutation. Obtaining data on this partitioning is difficult since it is not usually possible to measure the free energy associated with mutation in the native and denatured states. One study of a reversible disulfide mutant of T4 lysozyme indicated that the free energy of mutation operated mainly on the native state (Lu et al., 1992). In our case, because of the high degree of solvent exposure of lysine 73 (85−90% solvent-exposed) and its isolation from other charged side chains [>11 Å distance to the nearest charged residue; see Bowler et al. (1993)], we believe that if the native state structure is largely unperturbed, the majority of observed stability changes may be attributed to denatured state effects. Infrared data on the denatured state of iso-1-cytochrome *c* mutants with compact denatured states support this contention (Bowler et al., 1994). As noted above, some mutants in this study do show significant native state structural changes, and thus effects on native state energetics cannot be ignored in these cases.

Given this structural context, we will now examine the thermodynamics of the equilibrium between the native and denatured states for these mutants. First we will consider the control mutant, arginine 73 iso-1-cytochrome *c*. On the basis of charge, its thermodynamic parameters are expected to be very similar to the WT lysine 73 protein. Arginine shows a slight decrease in the midpoint of the transition from native to denatured states and a reasonably significant increase in the slope, *m*, of its $\Delta G^{\circ}_{u}$ versus [gdnHCl] plot. The overall effect is a similar value for $\Delta G^{\circ}_{n}{}^{H_2O}$. The change in the slope, *m*, merits some discussion as one might have expected nearly identical thermodynamic parameters for the lysine 73 and arginine 73 proteins. Some researchers have suggested that lysine to arginine replacements can stabilize the native state due to the increased hydrogen bonding potential of the guanidinium group (Mrabet et al., 1992). The high solvent exposure of lysine 73 likely precludes such an effect in this case. In a case similar to the Lys 73 to Arg variant, mutation of trimethylated lysine 72 of iso-1-cytochrome *c* to arginine (also highly solvent exposed) revealed a slight decrease in stability (Holzschu et al., 1987). Thus, highly solvent-exposed sites do not appear to show stabilizing effects due to the higher hydrogen bonding potential of arginine versus lysine. To understand the less compact denatured state of Arg 73 iso-1-cytochrome *c*, a strict hydrophilic/hydrophobic classification of these residues may not be adequate. Environmental characteristics of amino acids derived from the 3D MHP data on 23 protein structures indicated arginine residues exhibited significant nonhydrophilic properties, tending to be in contact with nonpolar residues (Efremov & Alix, 1993). In addition, correlation of hydrophobicity patterns of sequence sets with solvent accessibility patterns of known structures revealed that the long, aliphatic parts of arginine and lysine side chains can effectively substitute for hydrophobic amino acids at some buried positions and still permit the charged groups at the ends of the side chains to reach the solvent (Bowie et al., 1990). The arginine side chain has a smaller aliphatic component than lysine (three versus four methylenes), and the polar component of arginine has more hydrogen bonding potential and is more basic than lysine. Upon unfolding, arginine may seek a more solvated environment than lysine, leading to a more expanded denatured state.

The heteropolymer theory of Dill and co-workers, applied to single-site mutations (Shortle et al., 1992), indicates that hydrophilic to hydrophobic mutations at surface sites in the native state would be expected to stabilize the denatured state of a protein through increased hydrophobic interactions. The results of our first study were consistent with this notion, showing a correlation between the hydrophobicity of the residue replacing lysine with the decrease in the $\Delta G^{\circ}_{u}{}^{H_2O}$ for that mutant [$r = 0.98$, $P_N = 0.24\%$; see Bowler et al. (1993)]. In Figure 4, a plot of $\Delta G^{\circ}_{u}{}^{H_2O}$ versus 1-octanol to water transfer free energies, $\Delta G_{tr}$ (Fauchere & Pliska, 1983;

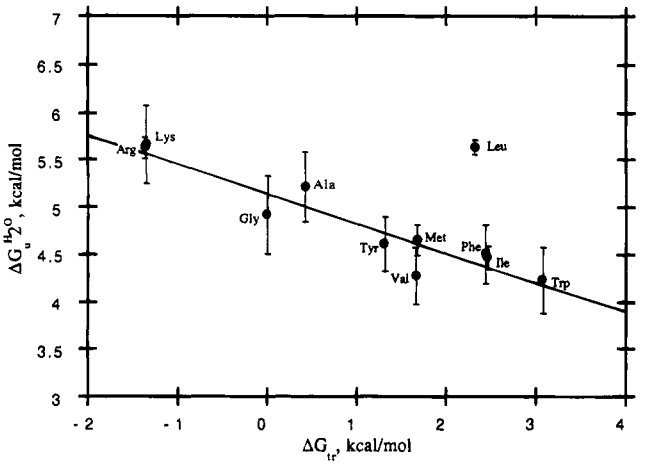


FIGURE 4: Plot of $\Delta G^\circ_{u}{}^{H_2O}$ for each mutant versus the hydrophobicity of the residue at position 73. The hydrophobicity, $\Delta G_{tr}$, is the free energy of transfer of the residue from 1-octanol to water (Eisenberg & McLachlan, 1986; Fauchere & Pliska, 1983). Error bars are one standard deviation for each $\Delta G^\circ_{u}{}^{H_2O}$ value, and a linear least-squares fit to the data is shown. The equation of this fit is $\Delta G^\circ_{u}{}^{H_2O} = -0.32(\Delta G_{tr}) + 5.14$, $r = 0.94$.

Eisenberg & McLachlan, 1986), is shown which includes data from our previous study (Bowler et al., 1993) as well as the six new data points from this work. A linear correlation is shown for 10 of the 11 data points in Figure 4. The correlation coefficient, $r$, is 0.94. The probability, $P_N$, that this is a chance correlation is 0.0053%, a highly significant correlation (Taylor, 1982). Inclusion of the outlier point, leucine, decreases $r$ to 0.71 ($P_N = 1.4\%$). A similar correlation can be obtained between the slope, $m$, and $\Delta G_{tr}$ with $r = 0.83$ and $P_N = 0.30\%$, again a highly significant correlation (for inclusion of leucine, $r = 0.67$, $P_N = 2.4\%$). The observed decrease in $\Delta G^\circ_{u}{}^{H_2O}$ with increasing hydrophobicity is what would be expected for a hydrophobic effect acting on the denatured state, whereas an increase in $\Delta G^\circ_{u}{}^{H_2O}$ would be expected for a hydrophobic effect acting on the native state. The correlation between the slope, $m$, and $\Delta G_{tr}$ suggests that the decrease in $\Delta G^\circ_{u}{}^{H_2O}$ is due to stabilization of the denatured state caused by enhancement of hydrophobic clusters in a more compact denatured state (*vide infra*). Thus, the hydrophobic to hydrophilic surface mutations at position 73 of iso-1-cytochrome *c* appear to act largely in accord with the predictions of heteropolymer theory; stabilization of the denatured state is achieved through increased hydrophobic interactions. The slope of the correlation in Figure 4, −0.32, is identical to that observed previously (Bowler et al., 1993) for the polarizable side chains (Trp, Tyr, Phe, Met) alone. Thus, our data do not provide any compelling evidence that the polarizability of a hydrophobic residue affects the favorability of hydrophobic cluster formation in the denatured state.

In general, two factors are expected to dominate the free energy of the denatured state: the entropy of the polypeptide chain and the solvation free energy of the nonpolar surface, also mainly an entropic term. These two forces oppose each other. In the limit of an expanded denatured state, the chain entropy term might be expected to dominate since hydrophobic interactions will be minimal. This appears to be the case for glycine to alanine and alanine to proline or valine mutants of T4 lysozyme (Matthews et al., 1987; S. Dao-pin et al., 1990). For proteins which appear to unfold to more compact denatured states, mutations affecting the compactness of the denatured state behave as though they are dominated by the solvation free energy term. More compact denatured states are more stable and lead to a decrease in $\Delta G^\circ_{u}{}^{H_2O}$ (Shortle et al., 1992). For staphylococcal nuclease, a decrease in $A_d$ for mutants with decreased slopes, $m$, with respect to wild-type protein ($m^-$ mutants) is related to a greater decrease in solvation free energy which outweighs the positive contribution to free energy due to chain entropy (Green et al., 1992). Such effects seem to be operative for iso-1-cytochrome *c* both in our work [this study and Bowler et al. (1993)] and in that of others (Betz & Pielak, 1992).

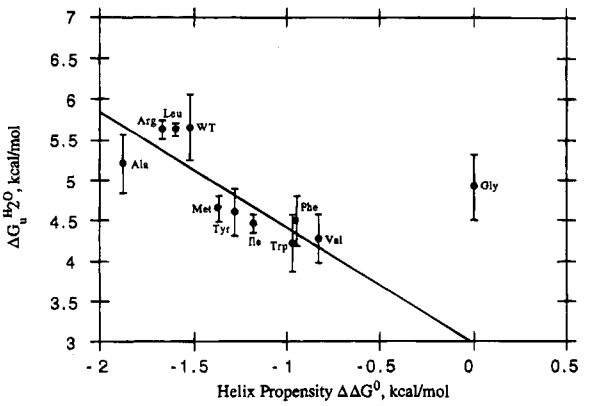


FIGURE 5: Plot of $\Delta G^\circ_{u}{}^{H_2O}$ versus α-helix propensity, $\Delta\Delta G^\circ$, for each mutant at position 73. The helix propensity of each amino acid is derived from experiments on alanine-based peptides, analyzed by modified Lifson−Roig theory (Chakrabartty et al., 1994). The equation of the linear regression fit is $\Delta G^\circ_{u}{}^{H_2O} = -1.42(\Delta\Delta G^\circ) + 3.01$, $r = 0.84$.

Given the perturbations to the native state structure for two of our mutants (Ile, Gly), we need to consider possible native state effects for this set of mutations. Lysine 73 occurs in a short stretch (one turn) of α-helix, and therefore it is reasonable to consider the effect of α-helix propensity on the set of mutations we have studied. Given the high solvent exposure of this site, it is appropriate to consider a correlation with Baldwin's α-helix propensity scale derived from studies on alanine-based α-helical peptides (Chakrabartty et al., 1994). It must be noted that Baldwin's scale is for midhelix residues. Since the α-helix under consideration only runs from positions 70 to 75 of iso-1-cytochrome *c*, both N-capping and C-capping effects would be expected for mutations to lysine 73. Thus, interpretation of this correlation may not be straightforward. Figure 5 shows a plot of $\Delta G^\circ_{u}{}^{H_2O}$ versus α-helix propensity for the 11 amino acids we have studied at position 73. The correlation shown is for all points except glycine 73 and gives $r = 0.84$ and $P_N = 0.24\%$, a correlation coefficient similar to that observed for other protein systems (Chakrabartty et al., 1994; Pinker et al., 1993; Blaber et al., 1993, 1994; Horovitz et al., 1992). This correlation is less significant than the one with $\Delta G_{tr}$, and inclusion of glycine decreases $r$ to 0.52 and $P_N$ to 10.1%, no longer a significant correlation. Still, these data suggest that α-helix propensity may be affecting the observed $\Delta G^\circ_{u}{}^{H_2O}$ values. We note that leucine 73 lies close to the correlation line shown in Figure 5, and thus α-helix propensity effects may explain some of the unusual properties of leucine with respect to the correlation in Figure 4. Correlations with other α-helix scales are as follows: $r = 0.64$ (O'Neil & DeGrado, 1990); $r = 0.78$ (Lyu et al., 1990); T4 lysozyme site 131, $r = 0.03$; site 44, $r = 0.25$ (Blaber et al., 1993, 1994); barnase $r = 0.77$ (Horovitz et al., 1992).

Correlation with factors believed to stabilize α-helices is as follows: loss of side chain entropy, Monte Carlo scale, $r = 0.29$ (Creamer & Rose, 1992, 1994); protein database scale, $r = 0.25$ (Blaber et al., 1994); and hydrophobic burial, $r = 0.07$ (Blaber et al., 1994). It is perhaps not surprising to find such poor correlations with loss of side chain entropy and hydrophobic burial since both these factors depend mainly on interactions with the previous turn of the α-helix which is only partially present ($i - 3$ is present, but $i - 4$ is not) for the one-turn helix which contains lysine 73. Overall, these correlations suggest a weak to moderate correlation with α-helix propensity.

It can be seen in Figures 4 and 5 that neither the correlation to $\Delta G_{tr}$ nor the correlation to α-helix propensity gives a slope of 1, which might be expected if either correlation accounted for all of the deviation in $\Delta G^{\circ}_{u}{}^{H_2O}$ upon mutation (slope = −1.42 for the α-helix propensity correlation and −0.32 for the $\Delta G_{tr}$ correlation). The variation in $\Delta G^{\circ}_{u}{}^{H_2O}$ is greater than would be expected if native state α-helix effects were the sole factor, and the variation in $\Delta G^{\circ}_{u}{}^{H_2O}$ should be larger if hydrophobicity effects in the denatured state acted to their full potential. The exact proportioning of these two effects is difficult to determine especially since the application of midhelix parameters to our system may be suspect. We have attempted a bilinear fit of our $\Delta\Delta G^{\circ}_{u}{}^{H_2O}$ values [where $\Delta\Delta G^{\circ}_{u}{}^{H_2O} = \Delta G^{\circ}_{u}{}^{H_2O}(\text{mutant}) - \Delta G^{\circ}_{u}{}^{H_2O}(\text{glycine})$] to the equation:

$$\Delta\Delta G^{\circ}_{u}{}^{H_2O} = a\Delta G_{tr} + b\Delta\Delta G^{\circ}_{helix} + c$$

and obtain $a = -0.23$, $b = -0.43$, and $c = -0.29$. This fit suggests that the effects of hydrophobicity and α-helix propensity are roughly additive since the coefficients are similar to the slopes of the individual correlations of all 11 data points (−0.23 versus −0.25 and −0.42 versus −0.56). However, it is apparent for side chains with large residuals in this fit that additivity breaks down and one effect predominates. The $\Delta\Delta G^{\circ}_{u}{}^{H_2O}$ for the leucine 73 variant (residual 0.84 kcal) is entirely accounted for by α-helix propensity effects whereas the $\Delta\Delta G^{\circ}_{u}{}^{H_2O}$ for the valine 73 variant (residual 0.33 kcal) is accounted for entirely by a denatured state hydrophobicity effect. The higher correlation coefficient we observe for the $\Delta G^{\circ}_{u}{}^{H_2O}$ versus $\Delta G_{tr}$ plot suggests that the denatured state stability effect for the position 73 mutants of iso-1-cytochrome *c* is more important, but is modulated by α-helix propensity effects. Studies on myoglobin mutants have also demonstrated the interaction of α-helix effects and denatured state compactness effects (*m* values) in determining the overall stability of a protein (Lin et al., 1993).

Recent calorimetric studies on staphylococcal nuclease (Carra et al., 1994) suggest that mutationally-induced changes in *m* values derived from a two-state treatment of gdnHCl unfolding data may in fact reflect the increase or decrease in the population of a compact intermediate in the unfolding pathway due to mutation. Although cytochrome *c* does not have an obvious two-domain structure like staphylococcal nuclease, we cannot exclude such an interpretation with our current data. Thermal denaturation and calorimetric studies are currently underway with these cytochrome *c* variants which should determine if such an effect is operative in our system.

## REFERENCES


Alexandrescu, A. T., Abeygunawardana, C., & Shortle, D. (1994) *Biochemistry 33*, 1063−1072.

Bennetzen, J. L., & Hall, B. D. (1982) *J. Biol. Chem. 257*, 3026−3031.

Betz, S. F., & Pielak, G. J. (1992) *Biochemistry 31*, 12337−12344.

Blaber, M., Zhang, X.-J., & Matthews, B. W. (1993) *Science 260*, 1637−1640.

Blaber, M., Zhang, X.-J., Lindstrom, J. D., Pepiot, S. D., Baase, W. A., & Matthews, B. W. (1994) *J. Mol. Biol. 235*, 600−624.

Bowie, J. U., Clarke, N. D., Pabo, C. O., & Sauer, R. T. (1990) *Proteins: Struct., Funct., Genet. 7*, 257−264.

Bowler, B. E., May, K., Zaragoza, T., York, P., Dong, A., & Caughey, W. S. (1993) *Biochemistry 32*, 183−190.

Bowler, B. E., Dong, A., & Caughey, W. S. (1994) *Biochemistry 33*, 2402−2408.

Carra, J. H., Anderson, E. A., & Privalov, P. (1994) *Biochemistry 33*, 10842−10850.

Chakrabartty, A., Kortemme, T., & Baldwin, R. L. (1994) *Protein Sci. 3*, 843−852.

Christianson, T. W., Sikorski, R. S., Dante, M., Shero, J. H., & Hieter, P. (1992) *Gene 110*, 119−122.

Creamer, T. P., & Rose, G. D. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 5937−5941.

Creamer, T. P., & Rose, G. D. (1994) *Proteins: Struct., Funct., Genet. 19*, 85−97.

Dao-pin, S., Baase, W. A., & Matthews, B. W. (1990) *Proteins: Struct., Funct., Genet. 7*, 198−204.

Dao-pin, S., Sauer, U., Nicholson, H., & Matthews, B. W. (1991a) *Biochemistry 30*, 7142−7153.

Dao-pin, S., Soderlind, E., Baase, W. A., Wozniak, J. A., Sauer, U., & Matthews, B. W. (1991b) *J. Mol. Biol. 221*, 873−887.

Dill, K. A., & Shortle, D. (1991) *Annu. Rev. Biochem. 60*, 795−825.

Dong, A., & Caughey, W. S. (1994) *Methods Enzymol. 232*, 139−175.

Dong, A., Huang, P., & Caughey, W. S. (1990) *Biochemistry 29*, 3303−3308.

Dong, A., Huang, P., & Caughey, W. S. (1992a) *Biochemistry 31*, 182−188.

Dong, A., Caughey, B., Caughey, W. S., Bhat, K. S., & Coe, J. E. (1992b) *Biochemistry 31*, 9364−9370.

Dong, A., Caughey, W. S., & Du Clos, T. W. (1994) *J. Biol. Chem. 269*, 6424−6430.

Dong, A., Haung, P., & Caughey, W. S. (1995) *Arch. Biochem. Biophys.* (in press).

Efremov, R. G., & Alix, A. J. P. (1993) *J. Biomol. Struct. Dyn. 11*, 483−507.

Eisenberg, D., & McLachlan, A. D. (1986) *Nature 319*, 199−203.

Fauchere, J. L., & Pliska, V. (1983) *Eur. J. Med. Chem. 18*, 369−375.

Faye, G., Leung, D. W., Tatchell, K., Hall, B. D., & Smith, M. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 2258−2262.

Flanagan, J. M., Kotaoka, M., Shortle, D., & Engelman, D. M. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 748−752.

Green, S. M., Meeker, A. K., & Shortle, D. (1992) *Biochemistry 31*, 5717−5728.

Hagihara, Y., Tan, Y., & Goto, Y. (1994) *J. Mol. Biol. 237*, 336−348.

Hickey, D. R., McLendon, G., & Sherman, F. (1988) *J. Biol. Chem. 263*, 18298−18305.

Horovitz, A., Matthews, J. M., & Fersht, A. R. (1992) *J. Mol. Biol. 227*, 560−568.

Ikemura, T. (1982) *J. Mol. Biol. 158*, 573−577.

James, E., Wu, P. G., Stites, W., & Brand, L. (1992) *Biochemistry 31*, 10217−10225.

Krimm, S., & Bandekar, J. (1986) *Adv. Protein Chem. 38*, 181−364.

Lin, L., Pinker, R. J., & Kallenbach, N. R. (1993) *Biochemistry 32*, 12638−12643.

Lu, J., Baase, W. A., Muchmore, D. C., & Dahlquist, F. W. (1992) *Biochemistry 31*, 7765−7772.

Lyu, P. C., Liff, M. I., Markey, L. A., & Kallenbach, N. R. (1990) *Science 250*, 669−673.

Matthews, B. W., Nicholson, H., & Becktel, W. J. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 6663−6667.
Mayr, L. M., & Schmid, F. X. (1993) *Biochemistry 32*, 7994−7998.
Montgomery, D. L., Hall, B. D., Gillam, S., & Smith, M. (1978) *Cell 14*, 673−680.
Mrabet, N. T., Van den Broeck, A., Van den brande, I., Stanssens, P., Laroche, Y., Lambeir, A.-M., Matthijssens, G., Jenkins, J., Chiadmi, M., van Tilbeurgh, H., Rey, F., Janin, J., Quax, W. J., Lasters, I., Demaeyer, M., & Wodak, S. J. (1992) *Biochemistry 31*, 2239−2253.
Nall, B. T. (1990) in *Protein Folding* (Gierasch, L. M., & King, J., Eds.) pp 198−207, AAAS, Washington, D.C.
Neri, D., Billeter, M., Wider, G., & Wuthrich, K. (1992) *Science 257*, 1559−1563.
Nozaki, Y. (1972) *Methods Enzymol. 26*, 43−50.
O'Neil, K. T., & DeGrado, W. F. (1990) *Science 250*, 646−651.
Pace, C. N. (1986) *Methods Enzymol. 131*, 267−280.
Pakula, A. A., & Sauer, R. T. (1990) *Nature 344*, 363−364.
Pinker, R. J., Lin, L., Rose, G. D., & Kallenbach, N. R. (1993) *Protein Sci. 2*, 1099−1105.
Ramdas, L., & Nall, B. T. (1986) *Biochemistry 25*, 6959−6964.
Saab-Rincon, G., Froebe, C. L., & Matthews, C. R. (1993) *Biochemistry 32*, 13981−13990.
Schellman, J. A. (1978) *Biopolymers 17*, 1305−1322.
Schlereth, D. D., & Mantele, W. (1993) *Biochemistry 32*, 1118−1126.
Serrano, L., Sancho, J., Hirschberg, M., & Fersht, A. R. (1992) *J. Mol. Biol. 227*, 544−549.
Seshadri, S., Oberg, K. A., & Fink, A. L. (1994) *Biochemistry 33*, 1351−1355.
Sherman, F., Fink, G. R., & Hicks, J. B. (1986) *Methods in Yeast Genetics*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Shoemaker, D. P., Garland, C. W., & Steinfeld, J. I. (1974) *Expreriments in Physical Chemistry*, 3rd ed., pp 51−58, McGraw-Hill Book Co., New York.
Shortle, D. (1989) *J. Biol. Chem. 264*, 5315−5318.
Shortle, D. (1993) *Curr. Opin. Struct. Biol. 3*, 66−74.
Shortle, D., & Meeker, A. K. (1986) *Proteins: Struct., Funct., Genet. 1*, 81−89.
Shortle, D., & Abeyhunawardana, C. (1993) *Structure 1*, 121−134.
Shortle, D., Chan, H. S., & Dill, K. A. (1992) *Protein Sci. 1*, 201−215.
Sosnick, T. R., & Trewhella, J. (1992) *Biochemistry 31*, 8329−8335.
Surewicz, W. K., Mantsch, H. H., & Chapman, D. (1993) *Biochemistry 32*, 389−394.
Susi, H., & Byler, D. M. (1986) *Methods Enzymol. 130*, 290−311.
Tsong, T. Y. (1974) *J. Biol. Chem. 249*, 1988−1990.
Tsong, T. Y. (1975) *Biochemistry 14*, 1542−1547.
Taylor, J. R. (1982) *An Introduction to Error Analysis*, pp 169, 248−249, University Science Books, Mill Valley, CA.

BI941986Q