Construction and Characterization of Monomeric Tryptophan Repressor: A Model for an Early Intermediate in the Folding of a Dimeric Protein†

Xiao Shao,‡ Preston Hensley,§ and C. Robert Matthews*,‡

Department of Chemistry and Center for Biomolecular Structure and Function, The Pennsylvania State University, University Park, Pennsylvania 16802, and Department of Macromolecular Sciences, SmithKline Beecham Pharmaceuticals, King of Prussia, Pennsylvania 19406

Received April 4, 1997; Revised Manuscript Received June 6, 1997⊗

ABSTRACT: Tryptophan repressor (TR) from *Escherichia coli* is a homodimer whose highly helical subunits intertwine in a complex fashion. A monomeric version of Trp repressor has been constructed by introducing a pair of polar amino acids at the hydrophobic dimer interface. Analytical ultracentrifugation was used to show that the replacement of leucine at position 39 with glutamic acid results in a monomer/dimer equilibrium whose dissociation constant is 1.11×10^{-4} M at 25 °C and pH 7.6. Tryptophan fluorescence, both near- and far-UV circular dichroism, and NMR spectroscopies demonstrated that, at the micromolar concentrations where the monomer predominates, secondary and tertiary structure are present. Hydrophobic dye-binding experiments showed that nonpolar surface is accessible in the monomeric form. The urea-induced equilibrium unfolding of monomeric L39E TR was monitored by circular dichroism, fluorescence, and absorbance spectroscopies. Coincident transitions show that the urea denaturation process follows a simple two-state model involving monomeric native and unfolded forms. The free energy at standard state in the absence of denaturant was estimated to be 2.37 ± 0.15 kcal mol^{-1}, and the sensitivity of the unfolding transition to denaturant, the *m* value, was 0.86 ± 0.04 kcal mol^{-1} M(urea)$^{-1}$ at pH 7.6 and 25 °C. The thermal denaturation transition occurred over a broad temperature range, suggesting either that the enthalpy change is small or that intermediates may exist. Kinetic studies showed that both the refolding and unfolding of the monomer were complete in the mixing dead time of stopped-flow CD and fluorescence spectroscopy, 5 ms. These structural, thermodynamic, and kinetic results are very similar to those previously reported for an early, monomeric intermediate in the folding of the wild-type TR dimer [Mann, C. J., & Matthews, C. R. (1993) *Biochemistry 32,* 5282−5290]. The construction of a stable, monomeric form of TR that strongly resembles a transient folding intermediate should provide useful insights into the nature of the early events in the folding of TR.

The folding of multimeric proteins requires the coordinated development of secondary, tertiary, and quaternary structure. Although much has been learned about the formation of secondary and tertiary structure through studies of the folding of single-subunit proteins (Kim & Baldwin, 1990; Matthews, 1993), rather less is known about how these reactions are coupled with the association reactions that are a requisite aspect of quaternary structure formation. Some insight into the coordination of these events has been provided by previous folding studies on a variety of oligomeric proteins (Clark et al., 1997; Goldberg et al., 1990; Jaenicke, 1987, 1991; Phili et al., 1993). Early efforts used classical kinetic analysis and chemical cross-linking to probe the development of quaternary structure during the folding of multisubunit proteins (Jaenicke, 1987, 1991). The subsequent development and application of stopped-flow optical techniques permitted insight into the formation of secondary and tertiary structure and its relationship to the formation of quaternary structure. For example, partially folded monomeric intermediates have been observed early in the folding of dimeric β_2 subunit of tryptophan synthase (Goldberg et al., 1990) and brain-derived neurotrophic factor (BDNF) (Phili et al., 1993). These intermediates were compact and possessed secondary structure but no rigid tertiary structure. These features are characteristic of a molten globule species, an often-observed folding intermediate (Kuwajima, 1989; Ptitsyn, 1996; Ptitsyn et al., 1990). However, detailed structural and thermodynamic characterization of these intermediates is made difficult by their short lifetimes.

One approach for probing elementary events in the folding of oligomeric proteins has been to use protein engineering to create stable, monomeric species that are incapable of associating to form the native, multisubunit complex. Well-folded monomeric λ repressor, λ_{6-85} was constructed by eliminating the dimerization interface and the first five amino acids in the wild-type sequence (Huang & Oas, 1995a; Jordan & Pabo, 1988; Weiss et al., 1984). A well-structured, monomeric version of the cro repressor was created by replacing an intermolecular β-strand/β-strand interaction with an intramolecular interaction. The success of this design was demonstrated in the crystal structure of this variant (Albright et al., 1996). Alanine substitutions at the dimeric interface of Arc repressor dimer resulted in a monomeric protein that, unlike its λ and cro counterparts, remains in a physiologically denatured state (Milla & Sauer, 1995). The relationship of these monomers to transient intermediates that might arise

† This work was supported by the National Institutes of Health through Grants GM23303 and GM54836 to C.R.M.

* Author to whom correspondence should be addressed.
‡ The Pennsylvania State University.
§ SmithKline Beecham Pharmaceuticals.

⊗ Abstract published in *Advance ACS Abstracts,* August 1, 1997.

S0006-2960(97)00778-2 CCC: $14.00

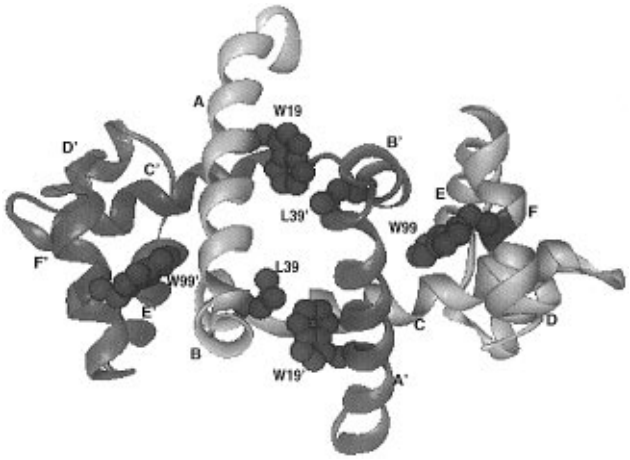

FIGURE 1: Ribbon diagram of the dimeric TR (Zhang et al., 1987), in which the individual subunits are shown in dark and light shading, respectively, and residues Trp 19, Trp 99, and Leu 39 are highlighted. The six helices in each monomer, A−F, are labeled.

during the folding of the wild-type dimeric species, unfortunately, was not addressed.

Trp repressor, another small, dimeric DNA binding protein, has been the subject of extensive kinetic folding studies (Gittelman & Matthews, 1990; Mann & Matthews, 1993; Mann et al., 1995) and offers an interesting system to test the relationship between a stable monomeric form and an early folding intermediate. Trp repressor is a highly helical, homodimer with 108 amino acids encoded in the DNA sequence for each monomer (Gunsalus & Yanofsky, 1980); posttranslational processing removes the N-terminal methionine residue (Joachimiak et al., 1983). Each subunit consists of six helices that are intertwined in a complex fashion (Figure 1) (Zhang et al., 1987). Helices A, B, and C from each monomer form the hydrophobic core and a principal part of the dimer interface. Helices D and E make up the helix−turn−helix DNA reading heads. Helix F docks against the core and contributes a part of the subunit interface. Previous studies have shown that the equilibrium unfolding of WT TR[1] follows a cooperative, two-state model, in which only native dimer and unfolded monomer are highly populated (Gittelman & Matthews, 1990). However, early in the kinetic folding process, a transient monomeric intermediate was proposed to appear (Mann & Matthews, 1993). This intermediate has significant secondary structure, a hydrophobic surface, and a stability of ∼3 kcal mol^{-1}. Further studies of its properties are hindered by a lifetime of less than several hundred milliseconds.

A clue as to how to create a stable monomer Trp repressor was provided by a previous protein engineering study in which each of the two tryptophan residues was individually replaced by phenylalanine (Mann et al., 1993). The noncoincidence of tryptophan fluorescence and far-UV CD urea denaturation curves in both of these mutants was interpreted in terms of the population of a stable, folded, monomeric intermediate. These results were rationalized in terms of the selective disruption of the hydrophobic environment at two very different part of the subunit interface where the two tryptophans reside (Figure 1). Because these intermediates only appeared in the presence of several molar urea and never exceeded 50% of the total population (Mann et al., 1993), they were not readily studied.

Reasoning that polar replacements in the hydrophobic core between the subunits might accentuate the subunit dissociation reaction without unfolding an individual subunit, a buried leucine was replaced with a glutamic acid. The construction and characterization of L39E, a mutant TR that is predominantly monomeric at micromolar concentrations and in the absence of denaturant, is described in this report.

MATERIALS AND METHODS

Mutagenesis and Protein Purification. L39E TR was constructed using PCR techniques (Dulau et al., 1989; Landt et al., 1990) to introduce a GAA codon at the desired position in the gene for wild-type TR. The wild-type gene had been previously cloned into a pT_7RUC plasmid containing the T7 promoter (Gloss & Matthews, 1997). The plasmid containing the L39E mutation was transformed into *Escherichia coli* strain BL21(DE3) (Studier et al., 1990) for expression. At mid-log phase, protein production was induced by the addition of 0.5 mM IPTG.

The cells harvested from 10 L of LB culture were resuspended in 50 mL of Tris lysis buffer (100 mM Tris·HCl, pH 7.6, and 20 mM $MgCl_2$) containing 0.5 mg mL^{-1} lysozyme and incubated for 30 min at room temperature. After disruption of the cells by repeated sonication, L39E TR was found to exist primarily in the supernatant of the resulting solution. After precipitation of polynucleic acids with 1% streptomycin sulfate and precipitation of undesired proteins with 20% ammonium sulfate, the supernatant was dialyzed against potassium phosphate buffer (10 mM potassium phosphate and 0.1 mM K_2EDTA, pH 7.6) at 4 °C. The dialysate was loaded onto a 1 L DE52 ion-exchange column (Pharmacia, Uppsala, Sweden) equilibrated in the same potassium phosphate buffer and eluted with a linear gradient of phosphate buffer containing 0−150 mM KCl at a flow rate of ∼1−2 mL min^{-1}. The protein was further purified on a Sephadex G-50 gel-filtration column (Pharmacia) in the above phosphate buffer containing 50 mM KCl. The purity of the final protein product was verified by the observation of a single band by $NaDodSO_4$−polyacrylamide gel electrophoresis. The purified protein was stored as a 70% ammonium sulfate precipitate and dialyzed against sodium phosphate buffer (10 mM sodium phosphate and 0.1 mM Na_2EDTA, pH 7.6) before use. The concentration of monomeric L39E TR was determined using the extinction coefficient of 1.4×10^4 M^{-1} cm^{-1} at 280 nm, calculated according to the method of Gill and von Hippel (1989).

Analytical Ultracentrifugation. Equilibrium ultracentrifuge data were collected on a Beckman XL-I analytical ultracentrifuge using double sector cells with sapphire windows (Hensley, 1996). The test for equilibrium was the superposition of curves that were collected 4 h apart. The data were fit to the sum of two exponentials, one for monomer and one for dimer. The primary data, absorbance at 280 nm versus radius, were fit to the following equation:

[1] Abbreviations: ANS, 8-anilino-1-naphthalenesulfonate; CD, circular dichroism; L39E TR, tryptophan aporepressor mutant in which leucine at position 39 is replaced by glutamic acid; Na_2EDTA, ethlylenediaminetetraacetic acid, disodium salt; $NaDodSO_4$, sodium dodecyl sulfate; FPLC, fast protein liquid chromatography; MALDI-MS, matrix-assisted laser desorption−ionization mass spectroscopy; WT TR, wild-type tryptophan aporepressor.

$$A_{tot} = c_m\epsilon_m \exp\left[\frac{M_m(1-\bar{\nu}_m\rho)\omega^2(r^2 - r_m^2)}{2RT}\right] + \frac{c_m^2}{K_{1,2}}(2\epsilon_m)\exp\left[\frac{M_d(1-\bar{\nu}_d\rho)\omega^2(r^2 - r_m^2)}{2RT}\right] \quad (1)$$

where A_{tot} is the total absorbance, c_m and ϵ_m are the concentrations of the monomer at a reference position and the molar extinction coefficient of the monomer, M and $\bar{\nu}$ are the molecular mass and partial specific volume, ρ is the solvent density, ω is the angular velocity, r and r_m are the measured and reference radial positions, $K_{1,2}$ is the dimer equilibrium dissociation constant, R is the gas constant, and T is the absolute temperature. The subscripts m and d for c, ϵ, M and $\bar{\nu}$ refer to monomer and dimer, respectively. The partial specific volume was estimated by the method of Cohn and Edsall to be 0.741 mL g^{-1}, and the solvent density was estimated to be 1.0017 g mL^{-1} (Laue et al., 1992). The molecular mass of the monomer was determined by MALDI-MS to be 12 240 Da. This value compares well with the molecular weight predicted from the sequence from residues 2 to 108, 12 225 Da; the amino-terminal methionine is known to be removed by proteolysis *in vivo* (Joachimiak et al., 1983).

FPLC Size-Exclusion Chromatography. A Superdex 75 HR 10/30 size-exclusion column (Pharmacia) was fitted to an FPLC system (Pharmacia) and equilibrated with sodium phosphate buffer (10 mM sodium phosphate and 0.1 mM Na_2EDTA, pH 7.6) containing 50 mM NaCl at a flow rate of 0.5 mL min^{-1}. The absorbance of the eluent was monitored at 280 nm. The molecular weight standards (Sigma, St. Louis, MO) used for calibration were monomeric bovine serum albumin (66 kDa), chicken egg albumin (45 kDa), ovalbumin (43 kDa), carbonic anhydrase (29 kDa), α-lactalbumin (14.2 kDa), and vitamin B12 (1.3 kDa).

Optical Spectroscopic Methods. CD spectra were collected on an AVIV Associates Model 62 DS circular dichroism spectrometer. For far-UV CD, the path length of the cuvette was either 1 or 2 mm. A 10 cm path length was used for near-UV CD experiments. The near-UV CD spectrum represented the average of 10 repeat scans. The UV absorbance was measured on an Aviv/Cary 118 DS spectrophotometer. Steady-state fluorescence emission spectra were taken on a Spex Fluorolog 1.81m spectrometer. Intrinsic tryptophan fluorescence emission was recorded from 310 to 450 nm with excitation at 295 nm. ANS fluorescence was collected from 450 to 600 nm with excitation at 370 nm. ANS concentration was determined using an extinction coefficient of 6800 M^{-1} cm^{-1} at 370 nm (Haugland, 1991).

Stopped-flow CD and fluorescence experiments were conducted on Bio-logic SFM-3 stopped-flow systems attached to CD and fluorescence spectrometers. The instrumental set-up and procedures have been previously described (Mann et al., 1995). The dead time of mixing on both instruments was determined to be 5 ms (Tonomura et al., 1978).

NMR Spectroscopy. The L39E TR protein precipitated in 70% ammonium sulfate was centrifuged and dissolved in a buffer containing 10 mM sodium phosphate, pH 7.6. The protein was dialyzed against this buffer and then lyophilized. To exchange the amide protons for deuterium, the lyophilized protein was incubated in ^2H_2O at 45 °C for about 60 min. The protein was then lyophilized twice, each time after being redissolved in ^2H_2O. The final NMR sample, which was centrifuged to remove any aggregates, contained 10 μM L39E TR in 10 mM sodium phosphate at p^{2}H 7.6. The p^{2}H of the sample was adjusted with ^{2}HCl and NaO^{2}H. No correction was made for the isotope effect on the glass electrode.

The NMR spectrum of L39E TR was recorded at 500 MHz on a Bruker AMX-2 500 spectrometer at 25 °C. The spectral width was 6048 Hz and was centered on the water resonance, 4.76 ppm. The water signal was suppressed by saturation with the decoupler prior to data acquisition. The spectrum is the average of 11 200 transients.

Urea and Thermal Denaturation. Ultrapure urea was purchased from ICN Biomedicals, Inc. (Costa Mesa, CA), and used without further purification. All other chemicals were reagent-grade. The buffer used in all experiments was 10 mM sodium phosphate and 0.1 mM Na_2EDTA at pH 7.6. Before collection of urea-induced unfolding data, protein samples were incubated at 25 °C for 1 h to ensure full equilibration. Identical scans at longer times confirmed the validity of this procedure. The thermal denaturation experiment was monitored by CD at 222 nm using a 1 °C step size, a 30 s averaging time, and 2 min of equilibration time at each temperature. The complete equilibration at a given temperature was demonstrated by the absence of further changes in signals at longer times.

Data Analysis. CD data were corrected to mean residue ellipticity. The fluorescence data were converted to center of mass wavelength for each spectrum using the function

$$\bar{\lambda} = \frac{\sum(\lambda_i I_i)}{\sum I_i} \quad (2)$$

where λ_i is the wavelength and I_i is the intensity at each wavelength.

The equilibrium unfolding data were fit to a two-state model:

$$N \overset{K}{\rightleftharpoons} U \quad (3)$$

where N is the native monomer, U is the unfolded monomer, and $K = [U]/[N]$. The apparent fraction of unfolded protein, F_{app}, is

$$F_{app} = K/(1 + K) \quad (4)$$

where $K = \exp(-\Delta G^o/RT)$. The apparent free energy difference between native and unfolded species, ΔG^o, was assumed to depend linearly on the denaturant concentration, $\Delta G^o = \Delta G^o(H_2O) - m[\text{urea}]$, where $\Delta G^o(H_2O)$ is the free energy difference in the absence of denaturant and m describes the sensitivity of the unfolding transition to urea (Matthews, 1987; Pace, 1986; Schellman, 1978).

For comparison of unfolding transitions collected at different protein concentrations and by different spectral techniques, data were converted to an apparent fraction of unfolded protein, F_{app}:

$$F_{app} = \frac{Y_O - Y_N}{Y_U - Y_N} \quad (5)$$

where Y_O is the observed optical value at a given urea

concentration and Y_N and Y_U are the calculated values for the native and unfolded forms at the same denaturant concentration. Y_N and Y_U were assumed to depend linearly on denaturant concentration:

$$Y_N = Y_N(H_2O) + m_N[\text{urea}] \quad (6a)$$

$$Y_U = Y_U(H_2O) + m_U[\text{urea}] \quad (6b)$$

where m_N and m_U are the slopes of the pre- and posttransition regions for the native and unfolded forms, respectively (Santoro & Bolen, 1988).

Combining the above equations, a single comprehensive equation was used to fit the equilibrium unfolding data sets:

$$Y_{obs} = (Y_U(H_2O) + m_U[\text{urea}]) + \{\exp[-(\Delta G^o(H_2O) - m[\text{urea}])/RT]/(1 + \exp[-(\Delta G^o(H_2O) - m[\text{urea}])/RT])\} \times [(Y_U(H_2O) + m_U[\text{urea}]) - (Y_N(H_2O) + m_N[\text{urea}])] \quad (7)$$

The nonlinear least-squares analysis software, NLIN (SAS Institute, Inc., Cary, NC), was used to obtain fits of individual data sets. A global analysis of the data obtained from different spectroscopic methods was fit using the in-house software package Savuka 5.0, where $\Delta G^o(H_2O)$ and m were linked global parameters.

RESULTS

The complex topology of dimeric WT TR implies that the hydrophobic interface between the subunits is distributed throughout the structure (Zhang et al., 1987) (Figure 1). Although this situation would appear to make the selection of an appropriate mutational site difficult, the appearance of a monomeric intermediate by the replacement of either Trp 19 or Trp 99 with Phe (Mann et al., 1993) suggested that the subunit association reaction is highly cooperative. In other words, a sufficiently destabilizing replacement at almost any interface residue in the dimer might result in a monomeric form of TR. Leu 39 was chosen as the target for mutagenesis because its side chain is almost completely buried at the subunit interface in the structural core of TR. Replacement of this nonpolar side chain with the polar, charged side chain of glutamic acid should favor the dissociation of the subunits. The homodimeric nature of TR provides additional impetus to the destabilization event because both subunits carry the mutation.

Oligomeric State of L39E TR. Equilibrium sedimentation was used to determine the oligomeric state of L39E TR across a range of concentrations from 10^{-6} to 10^{-4} M. The equilibrium sedimentation profile (Figure 2, middle panel) was not well fit by a one-exponential equation (data not shown). A two-exponential equation, assuming a dynamic equilibrium between monomer and dimer, was required to give random residuals (Figure 2, upper panel). The dimer equilibrium dissociation constant, $K_{1,2}$, was estimated to be $(1.11 \pm 0.04) \times 10^{-4}$ M. The distribution of monomer and dimer species as a function of total protein concentration (in terms of monomer concentration) was calculated using this dissociation constant and is shown in Figure 2 (lower panel). The monomeric form of L39E is the dominant species in solution below 10^{-4} M; the dimeric form dominates above this concentration of protein. The oligomeric state of L39E TR was also tested by FPLC size-exclusion chromatography. At a concentration of ~25 μM, L39E eluted as a single peak whose molecular mass was estimated to be 13 kDa by using a set of calibrated standards (data not shown). This value is within error of that predicted on the basis of the amino acid sequence for residues 2−108, 12 225 Da, demonstrating that the monomeric form of L39E is the predominant species in solution under these conditions. Given that the ultracentrifugation results suggest that the dimer should represent about 25% of the population at this protein concentration, the observation of a single peak at slightly higher molecular weight than predicted may be a reflection of fast exchange between monomer and dimer (Shalongo et al., 1987, 1989, 1992). In the majority of the studies that follow, the protein concentration was kept in a range where the dimeric form of L39E represented less than 10% of the population so as to emphasize the properties of the monomeric form.

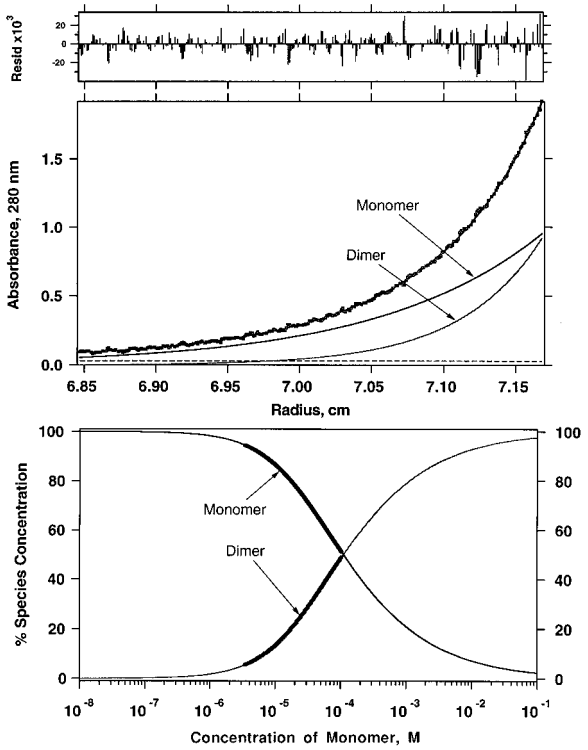

FIGURE 2: Equilibrium sedimentation data for L39E TR. (upper panel) Distribution of residuals between the fitted sum of two exponentials (one for monomer and one for dimer) and the observed data points. (middle panel) Fit of the primary data (absorbance vs radial position in centimeters) to eq 1 in Materials and Methods. Theoretical curves are indicated for monomer and dimer. The data from this single data set are best fit with $K_{1,2} = (1.11 \pm 0.04) \times 10^{-4}$ M. (lower panel) Calculated fractional distribution of monomer and dimer species as a function of total protein concentration in terms of monomer. The bold portion of the curve indicates the concentration range where the data were actually collected. The other portions of the curves are extrapolations based on the fitted parameter values. The protein was sedimented at 30 000 rpm for longer than 30 h at 25 °C. The buffer was 10 mM sodium phosphate and 0.1 mM Na_2EDTA at pH 7.6.

Secondary Structure of Monomeric L39E TR. A series of far-UV CD spectra of L39E TR at a series of increasing urea concentrations and at pH 7.6 and 25 °C are shown in Figure 3. These spectra were collected at 4.3 μM L39E TR, where the fraction of monomer is greater than 95%. In contrast to WT TR in the absence of urea, the relative magnitudes of the double minima at 208 and 222 nm indicate that both helical and random structure are present. The

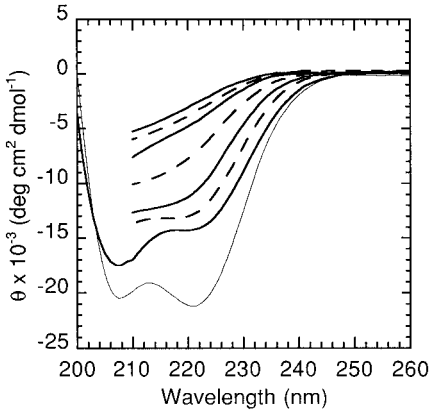

FIGURE 3: Far-UV CD spectra of L39E TR at different urea concentrations (thick lines) and WT TR at 0 M urea (thin solid line) at pH 7.6 and 25 °C. The alternating solid and dashed lines from bottom to top represent 0, 1, 2, 3, 4, 5, and 6 M urea, respectively. The concentration of L39E TR was 4.3 μM in the buffer described in the caption for Figure 2.

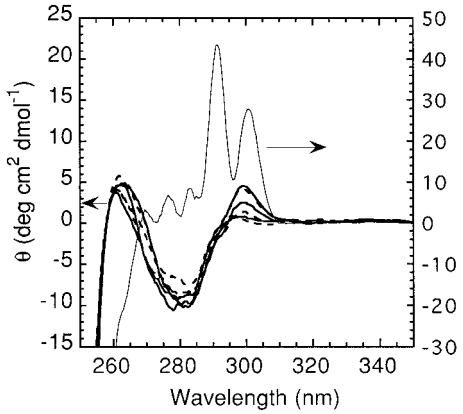

FIGURE 4: Near-UV CD spectra of L39E TR at different urea concentrations (thick lines) and WT TR at 0 M urea (thin solid line) at pH 7.6 and 25 °C. The alternating solid and dashed lines represent 0, 1, 2, 3, 4, and 6 M urea. At 300 nm, the signal progressively decreases. The protein concentration was 24.3 μM in the buffer described in the caption for Figure 2.

magnitude of the signal at 222 nm and at 0 M urea is about 67% that of WT TR (Mann et al., 1995), demonstrating that a significant fraction of sequence is helical. The progressive decrease in the signal as the urea concentration was increased shows that this secondary structure is disrupted at higher urea concentration, typical of globular proteins. To test for the possible influence of the dimer form, the urea denaturation was also performed at 19.3 μM protein. Within experimental error, the mean residue ellipticity at each urea concentration was unchanged even though the population of dimer species is predicted to be 18% at the higher protein concentration. The absence of an observable effect suggests that the dimer may be destabilized at urea concentrations required to unfold the monomeric form of L39E TR.

Tertiary Structure of Monomeric L39E TR. To test for the presence of aromatic side-chain packing in L39E TR, the near-UV CD spectrum was obtained at a series of urea concentrations (Figure 4). The presence of positive bands at 260 and 300 nm and a broad negative band between 270 and 290 nm in the absence of urea shows that unique packing environments exist for the aromatic side chains. The positive band at 300 nm disappears as the urea concentration is increased to 6 M, suggesting a loss of structure in an asymmetric environment near one or both of the tryptophan

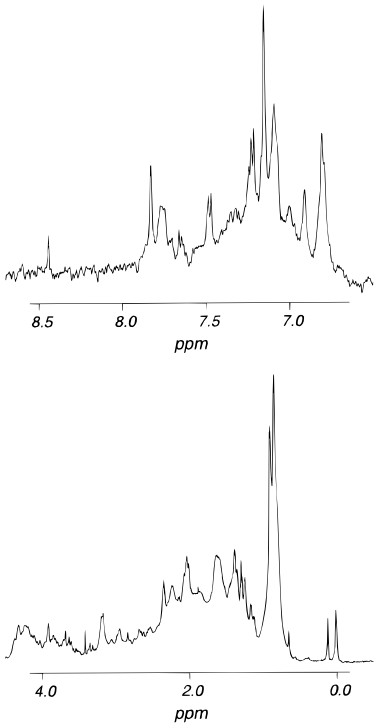

FIGURE 5: Aliphatic (−0.5 to 4.5 ppm) and aromatic (6.5 to 8.5 ppm) regions of the 1D proton NMR spectrum of monomeric L39E TR at 25 °C and p^{2}H 7.6. The protein concentration was 10 μM and the buffer was 10 mM sodium phosphate.

residues. Comparison of the spectra of folded L39E TR and WT TR (Figure 4) shows that the environments of the aromatic side chains are quite different in these two species.

The presence of well-defined side-chain packing in monomeric L39E TR was also assayed by 1D proton NMR spectroscopy. At 10 μM protein concentration and at 25 °C, p^{2}H 7.6, where the monomeric species is predicted to represent 86% of the population, several upfield-shifted resonances were observed near 0 ppm (Figure 5). This result and the dispersion observed in the aromatic region between 6.5 and 8.5 ppm (Figure 5) demonstrate that specific packing between aromatic and aliphatic side chains exists in the monomeric form of L39E TR.

The tertiary structure of L39E TR was further probed by collecting the intrinsic tryptophan fluorescence spectra at a protein concentration of 4.3 μM (Figure 6). In the absence of urea, the spectrum shows a maximum near 335 nm that is consistent with the burial of one or both tryptophans in a nonpolar environment. As the urea concentration was increased, the emission decreased in intensity and shifted to 350 nm, reflecting the exposure of one or both of the tryptophan side chains to solvent as the protein unfolds. The total red shift of L39E TR upon unfolding, represented by the center of mass wavelength value, is 10.0 nm, while the shift for WT TR is 16.5 nm (Mann et al., 1993). The smaller red shift for L39E indicates that the average environment of two tryptophans in the monomeric form is more polar than that in WT TR. Essentially the same results were obtained at 19.3 μM protein (data not shown), demonstrating either that the small fraction of dimeric form has little effect on the observed spectrum or thermodynamic properties or that urea disrupts the dimer.

A series of absorbance spectra of L39E TR at 24.3 μM protein and increasing urea concentrations were also collected

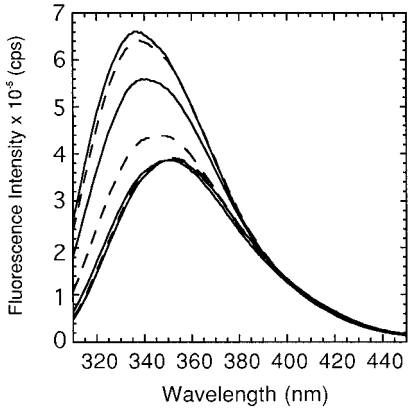

FIGURE 6: Intrinsic tryptophan emission spectra of L39E TR at different urea concentrations at pH 7.6 and 25 °C. The alternating solid and dashed lines from top to bottom represent 0, 1, 2, 3, 4, 5, and 6 M urea, respectively. The samples were excited at 295 nm. The protein concentration was 4.3 μM in the buffer described in the caption for Figure 2.

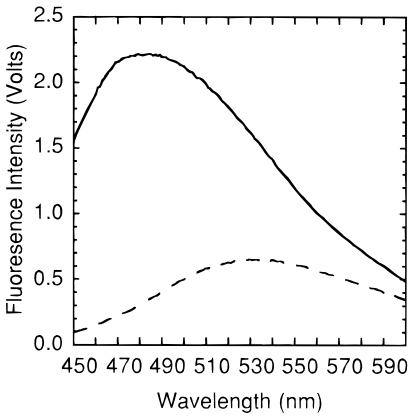

FIGURE 7: Representative spectra of free ANS (dashed line) and ANS bound to L39E TR (solid line) at 25 °C. Sample excitation was at 370 nm. The protein concentration was 5 μM and the ANS concentration was 11.9 μM in the buffer described in the caption for Figure 2.

(data not shown) to complement the near-UV CD data. The decrease in signal amplitude at 292 nm was consistent with the exposure of the tryptophan residues to solvent upon unfolding.

As a probe of accessible hydrophobic surface, 8-anilino-1-naphthalenesulfonate (ANS) was combined with L39E TR in the absence of denaturant. The fluorescence emission of ANS has been shown to undergo a significant blue shift and a dramatic increase in intensity when it binds to nonpolar surfaces on partially folded proteins (Ptitsyn et al., 1990; Tandon & Horowitz, 1989). The same behavior was observed with L39E TR (Figure 7), demonstrating that hydrophobic surfaces are present in the fully folded protein. To test whether the marginal stability of the monomeric mutant protein (see below) might permit the binding of ANS to alter its structure, the far-UV CD spectrum of L39E TR was obtained in the presence of excess ANS (data not shown). The absence of any effect of ANS on the spectrum demonstrated that ANS binding does not induce any significant reorganization of secondary structure. Further, a 25-fold excess of ANS did not alter the thermodynamic parameters describing the urea-induced unfolding transition of L39E TR (data not shown). Thus, the enhanced fluorescence of ANS in the presence of monomeric Trp repressor

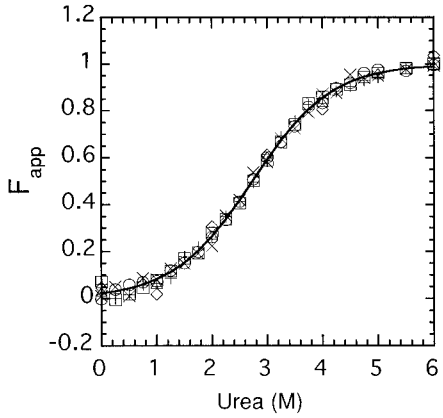

FIGURE 8: Dependence of the apparent fraction of unfolded protein, F_{app}, on the urea concentration at pH 7.6 and 25 °C as monitored by far-UV CD, tryptophan fluorescence, near-UV CD, and absorbance spectroscopies. The data shown in the plot were obtained from the far-UV CD signal at 222 nm at two protein concentrations, 4.3 μM (+) and 19.3 μM (×), the values of center of mass wavelength of fluorescence emission spectra at two protein concentrations, 4.3 μM (○) and 19.3 μM (□), and the near-UV signals at 300 nm (◇) and absorbance at 292 nm (△), both at a protein concentration of 24.3 μM. The line represents the global fit of the data to a two-state model.

accurately reflects the presence of nonpolar surfaces in its native conformation.

Stability of Monomeric L39E TR. The far- and near-UV CD and fluorescence data shown in Figures 3, 4, and 6 and the absorbance data (not shown) were normalized by conversion to apparent fraction of unfolded protein, F_{app}. The results are shown in Figure 8. The urea-induced equilibrium unfolding transitions measured by all of these optical techniques were fully reversible (data not shown). The coincidence of these transition curves shows that L39E undergoes a highly cooperative disruption of secondary and tertiary structure at increasing urea concentrations. Far-UV CD and fluorescence data sets collected at 4.3 and 19.3 μM gave identical transition curves, as did near-UV CD and absorbance data collected at 24.3 μM protein. As noted above, the observation that the variation of the predicted fraction of monomer from 75% to 95% had no effect on the apparent stability may reflect the effect of urea on the monomer/dimer equilibrium.

Fits of the individual data sets to a two-state model are summarized in Table 1, as are the results of a global fit across four different spectroscopic techniques and three different protein concentrations. The excellent agreement between the local and global fits strongly supports a two-state model for the urea-induced unfolding of L39E TR. The free energy of unfolding in the absence of denaturant is 2.37 ± 0.15 kcal mol^{-1}, the cooperativity parameter, m, is 0.86 ± 0.04 kcal mol^{-1} M^{-1}, and the midpoint of the transition is 2.77 ± 0.05 M urea (Table 1).

Temperature-Induced Denaturation. In an attempt to obtain other useful thermodynamic parameters, the thermal denaturation of L39E TR was also performed. A gradual decrease in ellipticity at 222 nm from -15.3×10^3 to -5.8×10^3 deg cm^2 dmol^{-1} was observed as the temperature was increased from 5 to 95 °C (Figure 9). The reversibility of the thermal denaturation was 95% (data not shown). The absence of definitive native and unfolded baselines made it impossible to fit these data to a model to determine the enthalpy, entropy, and heat capacity differences between

Table 1: Comparison of WT Dimeric TR, Monomeric L39E TR, and the Burst Phase Intermediate Species

	MRE_{222} at 0 M urea (10^3 deg cm^2 dmol^{-1})	$\Delta G^o(H_2O)$ (kcal mol^{-1})	m (kcal mol^{-1} M^{-1})	C_m (M)
WT dimeric TR[a]	21	22.4 ± 0.7	2.8 ± 0.1	5.6 ± 0.1
L39E monomeric TR				
4.3 μM (CD)		2.57 ± 0.39	0.96 ± 0.12	2.69 ± 0.12
4.3 μM (FL)		2.42 ± 0.35	0.90 ± 0.11	2.69 ± 0.11
19.3 μM (CD)		2.44 ± 0.55	0.89 ± 0.17	2.75 ± 0.17
19.3 μM (FL)		2.33 ± 0.22	0.83 ± 0.06	2.81 ± 0.09
global[b]	14	2.37 ± 0.15	0.86 ± 0.04	2.77 ± 0.05
burst phase species[a]	14	3.6 ± 0.3	1.0 ± 0.1	3.6

[a] The data were obtained from Mann and Matthews (1993). [b] Data sets used in the global fit: the far-UV signals at 222 nm at 4.3 and 19.3 μM; the values of center of mass wavelength of fluorescence at 4.3 and 19.3 μM; the near-UV signals at 300 nm, and absorbance at 292 nm at 24.3 μM. Conditions: 10 mM sodium phosphate and 0.1 mM Na_2EDTA, pH 7.6, and 25 °C.

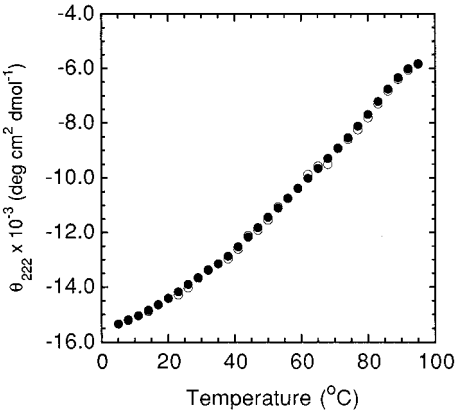

FIGURE 9: Thermal denaturation of L39E TR monitored by far-UV CD at 222 nm and at two different protein concentrations, 2.9 μM (○) and 10.8 μM (●).

these two states. Coincident temperature denaturation curves at 2.9 and 10.8 μM protein (Figure 9) imply that the observed behavior primarily reflects the thermodynamic properties of monomeric L39E TR.

Kinetic Folding Study. The time course of formation and disruption of secondary and tertiary structure in L39E TR was monitored by stopped-flow CD and fluorescence spectroscopy. For refolding and unfolding, the entire change in CD and tryptophan fluorescence intensity occurred within the dead time of mixing, 5 ms (data not shown). The complete formation of secondary and tertiary structure in less than 5 ms is similar to the behavior of other small globular proteins that lack transient intermediates (Huang & Oas, 1995b; Milla & Sauer, 1994; Robinson & Sauer, 1996).

DISCUSSION

Structural, Thermodynamic, and Kinetic Folding Properties of Monomeric L39E TR. The results of an equilibrium centrifugation study of L39E TR demonstrate that this mutant protein is primarily monomeric at micromolar concentrations. Spectroscopic and denaturation studies show that the monomeric form adopts a stable, well-structured conformation that resembles the native conformations of typical globular proteins in several respects. For example, far-UV CD, intrinsic tryptophan fluorescence, and absorbance spectroscopies revealed that L39E TR has substantial secondary structure and a hydrophobic core that is sufficient to at least partially bury one or both of its two tryptophans. Near-UV CD spectroscopy shows that one or both of its tryptophan side chains are constrained in a chiral environment, while proton NMR spectroscopy implies that specific hydrophobic interactions between aliphatic and aromatic side chains exist in the monomeric L39E TR. Finally, the two-state behavior of the urea-induced unfolding transition revealed that only a native and unfolded form are highly populated at equilibrium.

Other properties of monomeric L39E TR, however, suggest that its native conformation also resembles partially folded forms of other proteins. First, hydrophobic surface is readily accessible to ANS binding. This observation is consistent with an m value for the urea denaturation, 0.86 kcal mol^{-1} M(urea)$^{-1}$, that is smaller than that expected for a well-folded, 12.2 kDa protein, 1.3 kcal mol^{-1} M(urea)$^{-1}$ (Myers et al., 1995). The direct correlation between the exposure of buried nonpolar and polar surface to solvent accompanying unfolding and the magnitude of the m value implies that folded, monomeric L39E TR has more exposed nonpolar and polar surfaces than typical globular proteins of the same size. Second, the free energy of folding, 2.37 kcal mol^{-1}, is below the range observed for many other globular proteins, 5−15 kcal mol^{-1} (Pace et al., 1981). Third, the thermal denaturation profile exceeds the range from 0 to 95 °C, suggesting either that intermediates exist or that the enthalpy change for a two-state process is much smaller than that observed for most globular proteins (Freire & Murphy, 1991; Murphy & Freire, 1992; Privalov, 1989; Privalov & Potekhin, 1986).

A plausible explanation for these composite structural and thermodynamic properties in monomeric L39E TR can be obtained from inspection of the X-ray structure of WT TR (Figure 1). The hydrophobic core of the dimer is largely dominated by intersubunit nonpolar interactions. Because the dissociation of the two subunits would result in unfavorable interactions of these nonpolar surfaces with solvent, the polypeptide would undoubtedly collapse to a compact form. The intertwined topology of the two subunits implies that the hydrophobic core of this compact form, however, must differ substantially from the core of WT TR.

The hydrophobic core of the monomer displays elements of tertiary structure that differentiate it from the classical description of an unstructured molten globule (Kuwajima, 1989; Ptitsyn et al., 1990). When considered with its thermodynamic properties, the folded conformation of monomeric L39E TR might be best described as a structured molten globule (Kay & Baldwin, 1996; Ptitsyn, 1996). Highly ordered, partially folded states with specific, tight packing have also been observed in other proteins including apomyoglobin (Kay & Baldwin, 1996), cytochrome *c* (Mar-

morino & Pielak, 1995), and equine lysozyme (Morozova et al., 1995). The nativelike packing reported for these proteins, however, is unlikely to dominate the structure of monomeric L39E TR.

Comparison of Monomeric L39E TR with the Burst Phase Intermediate for WT TR. Previous kinetic studies on WT TR have shown that a cooperatively folded, monomeric burst phase species appears early in folding and precedes the dimeric association reaction (Mann & Matthews, 1993). Comparison of the properties of this transient intermediate with those of monomeric L39E TR shows several similarities: (1) The mean residue ellipticity at 222 nm of monomeric L39E TR under native conditions is 67% of the signal expected for the WT TR. This value is the same as the amplitude of the burst phase at 222 nm extrapolated to 0 M urea detected by stopped-flow CD (Mann & Matthews, 1993; Table 1). (2) Both monomeric L39E TR and the burst phase intermediate of WT TR unfold via a two-state mechanism. The stability of the monomeric L39E TR is comparable to, but somewhat less than, that of the burst phase intermediate, 2.37 versus 3.6 kcal mol^{-1}. The lower stability of L39E may reflect the effect of the mutation, suggesting that the side chain at position 39 is involved in structure in the burst phase intermediate for WT TR. (3) Comparable *m* values for the urea-induced unfolding of monomeric L39E TR and the burst phase species, 0.86 ± 0.04 and 1.0 ± 0.1 kcal mol^{-1} M(urea)$^{-1}$, respectively, imply that a similar amount of hydrophobic surface is exposed to solvent upon unfolding for both species. (4) The folded forms of both monomeric L39E TR and the burst phase intermediate contain exposed hydrophobic surfaces under native conditions, as monitored by ANS binding. (5) The folding of the L39E monomer and the burst phase intermediate are both complete within 5 ms.

Taken together, these spectroscopic, thermodynamic, and kinetic observations provide strong support for the conclusion that the monomeric folded form of L39E and the burst phase intermediate in the folding of WT TR are very similar species. Whether the probable nonnative packing in L39E TR and, by analogy, the burst phase intermediate helps or hinders the folding to the native, dimeric form must be tested in future experiments. Transient intermediates in other proteins (Kiefhaber et al., 1992; Weissman & Kim, 1995) have been found to retard the acquisition of the native conformation.

Equilibrium Folding Models for L39E TR and WT TR. If the monomer/dimer equilibrium of L39E TR is considered, the complete equilibrium unfolding model of L39E can be expressed as

$$N_2^* \rightleftharpoons 2F \rightleftharpoons 2U$$

where N_2* is the folded dimeric form, F is the folded monomeric form, and U is the unfolded form of L39E TR. The standard-state free energy change for the $N_2^* \rightleftharpoons 2F$ reaction is 5.4 kcal mol^{-1} (based on the dissociation constant of the L39E TR dimer, 1.11×10^{-4} M), while that for the $2F \rightleftharpoons 2U$ reaction is 4.74 kcal mol^{-1} (2×2.37 kcal mol^{-1}), respectively. Therefore, the total free energy difference between the native, dimeric form of L39E TR and its monomeric unfolded form is 10.1 kcal mol^{-1} at standard state.

The equilibrium unfolding process for WT TR has previously been shown to follow a two-state model involving only the native dimer and unfolded monomer (Gittelman & Matthews, 1990); the free energy of folding at standard state is 22.4 kcal mol^{-1} (Table 1). The magnitude of the destabilization obtained by replacing two leucines with two glutamic acid residues, 12.3 kcal mol^{-1}, is consistent with a previous observation by Stites et al. (1991). These investigators found that burying a single Lys side chain in the hydrophobic core of monomeric staphylococcal nuclease destabilizes the native conformation by 5.1 kcal mol^{-1}.

Relationship to Kinetic Folding Mechanism for WT TR. The existence of a stable, well-folded monomeric form for L39E TR that closely resembles a transient species detected during the folding of WT TR (Gittelman & Matthews, 1990; Mann & Matthews, 1993) should provide opportunities to obtain detailed thermodynamic, structural and dynamic information on an otherwise difficult-to-study form of TR. It would be especially interesting to learn what alternative side-chain packing arrangements are possible for a single TR subunit, prior to its self-association to form the native, homodimer structure. This insight should increase the understanding of the mechanism by which the amino acid sequence of Trp repressor is translated through a series of transient intermediates into its complex, three-dimensional structure.

ACKNOWLEDGMENT

We thank Dr. Dean E. McNulty for doing the mass spectral analysis and Peter J. Gualfetti for helping to carry out NMR experiments. We thank Drs. Jill Zitzewitz and Lisa Gloss for critical reviews of the manuscript. We are grateful to Drs. David Lambright and Osman Bilsel for the development of in-house global fitting software.

REFERENCES

Albright, R. A., Mossing, M. C., & Matthews, B. W. (1996) *Biochemistry 35*, 735−742.
Clark, A. C., Raso, S. W., Sinclair, J. F., Ziegler, M. M., Chaffotte, A. F., & Baldwin, T. O. (1997) *Biochemistry 36*, 1891−1899.
Dulau, L., Cheyrou, A., & Aigle, M. (1989) *Nucleic Acids Res. 17*, 2873.
Freire, E., & Murphy, K. P. (1991) *J. Mol. Biol. 222*, 687−698.
Gill, S. C., & von Hippel, P. H. (1989) *Anal. Biochem. 182*, 319−326.
Gittelman, M. S., & Matthews, C. R. (1990) *Biochemistry 29*, 7011−7020.
Gloss, L. M., & Matthews, C. R. (1997) *Biochemistry* (in press).
Goldberg, M. E., Semisotnov, G. V., Friguet, B., Kuwajima, K., Ptitsyn, O. B., & Sugai, S. (1990) *FEBS Lett. 263*, 51−56.
Gunsalus, R. P., & Yanofsky, C. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 7717−7721.
Haugland, R. P. (1991) *Handbook of Fluorescent Probes and Research Chemicals*, Catalog, Molecular Probes, Inc., Eugene, OR.
Hensley, P. (1996) *Structure 4*, 367−373.
Huang, G. S., & Oas, T. G. (1995a) *Biochemistry 34*, 3884−3892.
Huang, G. S., & Oas, T. G. (1995b) *Proc. Natl. Acad. Sci. U.S.A. 92*, 6878−6882.
Jaenicke, R. (1987) *Prog. Biophys. Mol. Biol. 49*, 117−237.
Jaenicke, R. (1991) *Biochemistry 30*, 3147−3161.
Joachimiak, A., Kelley, R. L., Gunsalus, R. P., Yanofsky, C., & Sigler, P. B. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 668−672.
Jordan, S. R., & Pabo, C. O. (1988) *Science 242*, 893−899.
Kay, M. S., & Baldwin, R. L. (1996) *Nat. Struct. Biol. 3*, 439−445.
Kiefhaber, T., Grunert, H.-P., Hahn, U., & Schmid, F. X. (1992) *Proteins: Struct., Funct., Genet. 12*, 171−179.
Kim, P. S., & Baldwin, R. L. (1990) *Annu. Rev. Biochem. 59*, 631−660.

Kuwajima, K. (1989) *Proteins: Struct., Funct., Genet. 6*, 87−103.
Landt, O., Grunert, H., -P., & Hahn, U. (1990) *Gene 96*, 125−128.
Laue, T. M., Shah, B. D., Ridgeway, T. M., & Pelletier, S. M. (1992) *Computer-Aided Interpretation of Analytical Sedimentation Data for Proteins* , Royal Society of Chemistry, Cambridge, U.K.
Mann, C. J., & Matthews, C. R. (1993) *Biochemistry 32*, 5282−5290.
Mann, C. J., Royer, C. A., & Matthews, C. R. (1993) *Protein Sci. 2*, 1853−1861.
Mann, C. J., Shao, X., & Matthews, C. R. (1995) *Biochemistry 34*, 14573−14580.
Marmorino, J. L., & Pielak, G. J. (1995) *Biochemistry 34*, 3140−3143.
Matthews, C. R. (1987) *Methods Enzymol. 154*, 498−511.
Matthews, C. R. (1993) *Annu. Rev. Biochem. 62*, 653−683.
Milla, M. E., & Sauer, R. T. (1994) *Biochemistry 33*, 1125−1133.
Milla, M. E., & Sauer, R. T. (1995) *Biochemistry 34*, 3344−3351.
Morozova, L. A., Haynie, D. T., AricoMuendel, C., Van, D. H., & Dobson, C. M. (1995) *Nat. Struct. Biol. 2*, 871−875.
Murphy, K. P., & Freire, E. (1992) *Adv. Protein Chem. 43*, 313−61.
Myers, J. K., Pace, C. N., & Scholtz, J. M. (1995) *Protein Sci. 4*, 2138−2148.
Pace, C. N. (1986) *Methods Enzymol. 131*, 266−280.
Pace, C. N., Fisher, L. M., & Cupo, J. F. (1981) *Acta Biol. Med. Ger. 40*, 1385−1392.
Phili, J. S., Rosenfeld, R., Arakawa, T., Wen, J., & Narhi, L. O. (1993) *Biochemistry 32*, 10812−10818.
Privalov, P. L. (1989) *Annu. Rev. Biophys. Biophys. Chem. 18*, 47−69.
Privalov, P. L., & Potekhin, S. A. (1986) *Methods Enzymol. 131*, 4−51.
Ptitsyn, O. (1996) *Nat. Struct. Biol. 3*, 488−490.
Ptitsyn, O. B., Pain, R. H., Semisotnov, G. V., Zerovnik, E., & Razgulyaev, O. I. (1990) *FEBS Lett. 262*, 20−24.
Ramakrishnan, V., & Gerchman, S. E. (1991) *J. Biol. Chem. 266*, 880−885.
Robinson, C. R., & Sauer, R. T. (1996) *Biochemistry 35*, 13878−13884.
Santoro, M. M., & Bolen, D. W. (1988) *Biochemistry 27*, 8063−8068.
Schellman, J. A. (1978) *Biopolymers 17*, 1305−1322.
Shalongo, W., Ledger, R., Jagannadham, M. V., & Stellwagen, E. (1987) *Biochemistry 26*, 3135−3141.
Shalongo, W., Jagannadham, M., Flynn, C., & Stellwagen, E. (1989) *Biochemistry 28*, 4820−4825.
Shalongo, W., Jagannadham, M., Heid, P., & Stellwagen, E. (1992) *Biochemistry 31*, 11390−11396.
Stites, W. E., Gittis, A. G., Lattman, E. E., & Shortle, D. (1991) *J. Mol. Biol. 222*, 7−14.
Tandon, S., & Horowitz, P. M. (1989) *J. Biol. Chem. 264*, 9859−9866.
Tonomura, B., Nakatani, H., Ohnishi, M., Yamaguchi-Ito, J., & Hiromi, K. (1978) *Anal. Biochem. 84*, 370−383.
Weiss, M. A., Sauer, R. T., Patel, D. J., & Karplus, M. (1984) *Biochemistry 23*, 5090−5095.
Weissman, J. S., & Kim, P. S. (1995) *Nat. Struct. Biol. 2*, 1123−1130.
Zhang, R.-G., Joachimiak, A., Lawson, C. L., Schevitz, R. W., Otwinowski, Z., & Sigler, P. B. (1987) *Nature 327*, 591−597.

BI9707786