Apparent Radii of the Native, Stable Intermediates and Unfolded Conformers of the α-Subunit of Tryptophan Synthase from E. coli, a TIM Barrel Protein †

ARTICLE *in* BIOCHEMISTRY · NOVEMBER 1999

Impact Factor: 3.02 · DOI: 10.1021/bi991296s · Source: PubMed

7 AUTHORS, INCLUDING:

Masahiro Iwakura
National Institute of Advanced Industrial Sci…
68 PUBLICATIONS **1,358** CITATIONS
SEE PROFILE

Hiroshi Kihara
Himeji-Hinomoto College
192 PUBLICATIONS **1,738** CITATIONS
SEE PROFILE

Jill A Zitzewitz
University of Massachusetts Medical School
21 PUBLICATIONS **997** CITATIONS
SEE PROFILE

Charles Robert Matthews
University of Massachusetts Medical School
146 PUBLICATIONS **5,902** CITATIONS
SEE PROFILE

Apparent Radii of the Native, Stable Intermediates and Unfolded Conformers of the α-Subunit of Tryptophan Synthase from *E. coli*, a TIM Barrel Protein†

Peter J. Gualfetti,‡,§ Masahiro Iwakura,|| J. Ching Lee,⊥ Hiroshi Kihara,# Osman Bilsel,‡ Jill A. Zitzewitz,‡ and C. Robert Matthews*,‡

Department of Chemistry and Center for Biomolecular Structure and Function, The Pennsylvania State University, University Park, Pennsylvania 16802, National Institute of Bioscience and Human Technology, Tsukuba, Ibaraki, Japan, Department of Human Biological Chemistry and Genetics, University of Texas Medical Branch, Galveston, TX 77550, Physics Laboratory, Kansai Medical University, Uyama-higashi, Hirakata, Osaka 573-0036 Japan

Received June 8, 1999; Revised Manuscript Received August 9, 1999

ABSTRACT: The urea-induced equilibrium unfolding of the α-subunit of tryptophan synthase (αTS) from *Escherichia coli* can be described by a four-state model, N ⇌ I1 ⇌ I2 ⇌ U, involving two highly populated intermediates, I1 and I2 [Gualfetti, P. J., Bilsel, O., and Matthews, C. R. (1999) *Protein Sci. 8*, 1623−1635]. To extend the physical characterization of these stable forms, the apparent radius was measured by several techniques. Size-exclusion chromatography (SEC), analytical ultracentrifugation (UC), and dynamic light scattering (DLS) experiments yield an apparent Stokes radius, R_s, of ∼24 Å for the native state of αTS. The small-angle X-ray scattering (SAXS) experiment yields a radius of gyration, R_g, of 19.1 Å, consistent with the value predicted from the X-ray structure and the Stokes radius. As the equilibrium is shifted to favor I1 at ∼3.2 M and I2 at 5.0 M urea, SEC and UC show that R_s increases from ∼38 to ∼52 Å. Measurements of the radius by DLS and SAXS between 2 and 4.5 M urea were complicated by the self-association of the I1 species at the relatively high concentrations required by those techniques. Above 6 M urea, SEC and UC reveal that R_s increases linearly with increasing urea concentration to ∼54 Å at 8 M urea. The measurements of R_s by DLS and R_g by SAXS are sufficiently imprecise that both values appear to be identical for the I2 and U states and, considering the errors, are in good agreement with the results from SEC and UC. Thermodynamic parameters extracted from the SEC data for the N ⇌ I1 and I1 ⇌ I2 transitions agree with those from the optical data, showing that this technique accurately monitors a part of the equilibrium model. The lack of sensitivity to the I2 ⇌ U transition, beyond a simple swelling of both species with increasing urea concentration, implies that the Stokes radii for the I2 and U states are not distinguishable. Surprisingly, the hydrophobic core known to stabilize I2 at 5.0 M urea [Saab-Rincón, G., Gualfetti, P. J., and Matthews, C. R. (1996) *Biochemistry 35*, 1988−1994] develops without a significant contraction of the polypeptide, i.e., beyond that experienced by the unfolded form at decreasing urea concentrations. Kratky plots of the SAXS data, however, reveal that I2, similar to N and I1, has a globular structure while U has a more random coil-like form. By contrast, the formation of substantial secondary structure and the burial of aromatic side chains in I1 and, eventually, N are accompanied by substantial decreases in their Stokes radii and, presumably, the size of their respective conformational ensembles.

A common characteristic of protein folding reactions is a dramatic decrease in the size of the conformational ensembles. While the native conformation usually reflects a sufficiently small ensemble that it can often be crystallized, the fully unfolded state appears to retain very little structure and can be well-described as a dynamic equilibrium between a vast number of conformers (*1*−*3*). Information on the size of these ensembles for stable species and for transient intermediates is essential in understanding the mechanism by which the amino acid sequence of a protein directs its folding to a native, well-defined conformation.

Measurements of global properties related to the size of these ensembles are provided by hydrodynamic methods (*4*) such as viscometry, ultracentrifugation, and size-exclusion chromatography, by scattering methods (*4*) such as light, X-ray, and neutron scattering and by fluorescence anisotropy measurements of rotational dynamics (*5*). Measurements of point-to-point distances by fluorescence resonance energy transfer complement and refine the global perspective (*6*, *7*).

For native states, the hydrodynamic techniques yield values for the Stokes radius, R_s,[1] and the scattering techniques yield

† This work was supported by the National Institute of Health through Grant GM 23303 to C.R.M.

* To whom correspondence should be addressed. Phone: (814) 865-8859. Fax (814) 863-8403. E-mail: crm@psu.edu.

‡ Department of Chemistry and Center for Biomolecular Structure and Function, The Pennsylvania State University, University Park, PA 16802.

§ Current address: Genencor International Inc., 925 Page Mill Road, Palo Alto, CA 94304.

|| National Institute of Bioscience and Human Technology, Tsukuba, Ibaraki, Japan.

⊥ Department of Human Biological Chemistry and Genetics, University of Texas Medical Branch, Galveston, TX 77550.

Physics Laboratory, Kansai Medical University, Uyama-higashi, Hirakata, Osaka 573-0036.

10.1021/bi991296s CCC: $18.00
Published on Web 09/11/1999

values for the radius of gyration, R_g, that usually agree well with those predicted from the X-ray crystal structure with or without a layer of bound water. Stabilized versions of molten globules, partially folded forms that are rich in secondary structure but lack defined tertiary structures (*8*−*11*), are found to have Stokes radii that are 10−30% larger than the native state (*12, 13*) but substantially smaller than the unfolded state ensemble. Because transient versions of molten globules often appear within the first few milliseconds of folding (*14, 15*), the contraction of conformational space available to the polypeptide could limit the search process and, potentially, enhance the folding reaction. Even in cases where such early intermediates act as kinetic traps, i.e., retard folding (*16, 17*), their significant influence on folding reactions justifies efforts to understand their physical properties. Finally, the size of an unfolded ensemble has been examined by Baskakov and Bolen (*18*) who used SEC to study the urea-induced unfolding of staphylococcal nuclease and several mutants. Consistent with previous arguments presented by Shortle (*19*), the size of the denatured state was found to vary significantly with the denaturant concentration.

A particularly interesting candidate for studies of the size of partially folded forms that play important roles in its folding mechanism is the α-subunit of tryptophan synthase from *Escherichia coli* (αTS). The three-dimensional structure of *E. coli* αTS is presumed to be very similar to the $(\beta/\alpha)_8$ motif observed in the crystal structure of the $\alpha_2\beta_2$ Trp synthase complex from *Salmonella typhimurium* (*20*). The α-subunit homologues are 85% identical in sequence and are each capable of activating both types of β_2 subunits (*21*). The TIM-barrel motif involves eight parallel strands, each of which hydrogen bonds to the preceding and succeeding strand in the sequence. Hydrogen bonds between strands 1 and 8 close the barrel (Figure 1). The β-barrel forms the core of the molecule with the interspersed amphipathic helices traversing the exterior of the barrel and serving to solubilize the protein.

Surprisingly, this single domain protein unfolds through a pair of stable intermediates when titrated with urea (*22, 23*).

$$N \leftrightarrows I1 \leftrightarrows I2 \leftrightarrows U \quad (1)$$

I1 is highly populated at ∼3.2 M urea and retains substantial fractions of the native secondary and tertiary structure. I2 is highly populated at ∼5.0 M urea and has little detectable secondary or tertiary structure. Fluorescence anisotropy and 1D proton NMR measurements on I2 indicate that a subset of tyrosine and histidine residues participate in a hydrophobic cluster (*24*, *25*). Kinetic studies have shown that I1 is an on-pathway intermediate whose interconversion to the native state, N, is one of the rate-limiting steps in folding (*17*). The role of I2 in the kinetic mechanism is not yet defined; however, available information suggests that I2 is an on-pathway species which appears in less than 5 ms after refolding from the unfolded state, U. The relatively rare opportunity to examine the size of one and possibly two on-pathway folding intermediates for αTS under equilibrium conditions motivated the present study.

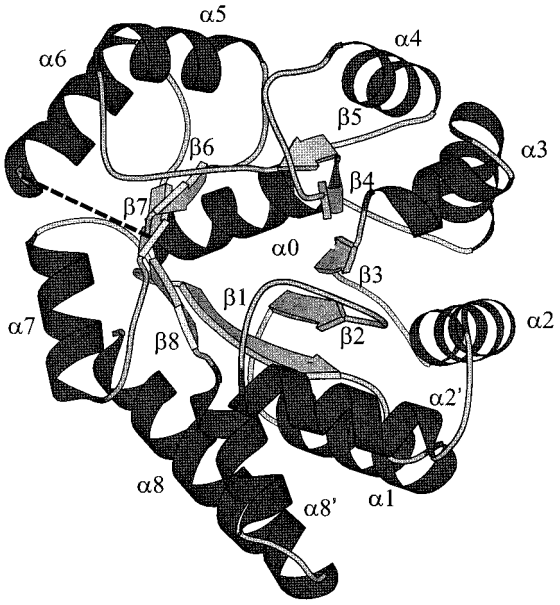

FIGURE 1: The X-ray crystal structure of the α-subunit of tryptophan synthase from *Salmonella typhimurium* (*20*) (personal communication from David Davies). The dashed line represents residues 178−192 that are undefined in the crystal structure.

MATERIALS AND METHODS

Chemicals and Reagents. Protein standards for the SEC calibration curve were purchased from Sigma at 95% purity or higher. All other chemicals were of reagent grade.

Protein Isolation and Purification. Wild-type αTS was expressed by *E. coli* strain W3110 (ΔtonB-trp)BA17*his*$^-$ containing the plasmid pBN55. The purification procedure was based on the method of Kirschner et al. (*26*) with modifications previously described by Gualfetti et al. (*24*). A purity of at least 95% was demonstrated by the observation of a single band on SDS−polyacrylamide gel electrophoretograms.

The presence of functional αTS was determined by the ability to activate the β_2 subunit of tryptophan synthase from *E. coli* (*27*). The concentration of the protein was determined by absorbance at 278 nm using an extinction coefficient of ∼12 700 cm^{-1} M^{-1} (*22*). The protein was stored as a precipitant in 70% saturated ammonium sulfate, 10 mM potassium phosphate, 4 mM K_2EDTA, and 1 mM DTE at pH 7.80 and 4 °C. The precipitated protein was prepared for experiments by resuspension in 10−100 mM potassium phosphate, 0.2 mM K_2EDTA and 1 mM DTE at pH 7.80 and dialysis against 4 × 1 L of the appropriate buffer at 4 °C.

Size-Exclusion Chromatography. Equilibrium SEC experiments were performed on a Pharmacia (Uppsala, Sweden) FPLC system with a UV-M II variable wavelength absorbance detector set at 280 nm. Experiments were run on a Superose 12 HR 10/30 column from Pharmacia modified

[1] Abbreviations: αTS, the α-subunit of tryptophan synthase from *E. coli*; BME, 2-mercaptoethanol; CD, circular dichroism; DLS, dynamic light scattering; DTE, dithioerythritol; FL, fluorescence; FPLC, fast-protein liquid chromatography; GdnHCl, guanidine hydrochloride; I1, an equilibrium intermediate of αTS maximally populated at 3.2 M urea; I2, an equilibrium intermediate of αTS maximally populated at 5.0 M urea; K_2EDTA, ethylenediaminetetraacetic acid, dipotassium salt; N, the native state; R_g, radius of gyration; R_s, Stokes radius; SAXS, small-angle X-ray scattering; SDS, sodium dodecyl sulfate; SEC, size-exclusion chromatography; SVD, singular value decomposition; TRA, time-resolved anisotropy; U, the unfolded state; UC, ultracentrifugation; V_{el}, elution volume.

in-house to include a water jacket for temperature regulation. The data were collected by a Dataq Instruments Inc. (Akron, Ohio) DI-130 digitizer supported by Windaq 100 software for Windows. The Stokes radius was determined by analysis of the elution volume with respect to a calibration curve prepared as previously described by Uversky (*13*). The 12 standard proteins used for the calibration curve were insulin (5.8 kDa for native and 2.9 kDa for unfolded), ubiquitin (8.5 kDa), cytochrome *c* (11.7 kDa), ribonuclease (13.7 kDa), lysozyme (14.3 kDa), hemoglobin (64.5 kDa for native and 15.5 kDa for unfolded), β-lactoglobulin (36.8 kDa for native and 18.4 kDa for unfolded), carbonic anhydrase (28.8 kDa), aldolase (160 kDa for native and 40.0 kDa for unfolded), bovine serum albumin (66.3 kDa), catalase (220 kDa) and ferritin (450 kDa). Samples were prepared by diluting 100 μL of a protein stock solution into 900 μL of premixed urea/buffer solutions. The final buffer conditions were 100 mM potassium phosphate, 0.2 mM K_2EDTA, and 1 mM DTE at pH 7.80, with various final concentrations of urea ranging 0−8 M. All samples were allowed to equilibrate at 25 °C for 1 h after which no further change in elution volume occurred. The urea concentration for each sample was calculated from its refractive index determined on a Reichert-Jung ABBE MARK II digital refractometer (*28*). Each sample was injected in triplicate in a volume of 100 μL and run at 0.7 mL min^{-1} after the column had been equilibrated with 1−2 column volumes of buffer or buffer/urea solutions. The protein concentration injected onto the column was approximately 10 μM.

The observation of Gaussian profiles at most urea concentrations permitted the V_{el} to be measured as the volume at the peak maximum. However, non-Gaussian profiles near 3 M urea, reflecting the relatively slow interchange of N and I1, required a more complex analysis. For those cases, the intensity-weighted average elution volume, or center-of-mass of the elution profile, was calculated using the following equation:

$$\langle V_{el,j} \rangle = (\sum A_{280,ij} V_{el,i})/(\sum A_{280,ij}) \quad (2)$$

where $A_{280,ij}$ represents the absorbance at 280 nm of the *i*th elution volume, $V_{el,i}$, at a particular urea concentration, *j*, and the summation is over the elution volume index, *i*. The assumption that N and I1 have equivalent extinction coefficients at 280 nm is justified by the observation that the extinction coefficients for the N and U forms at this wavelength differ by less than 15% (data not shown).

Ultracentrifugation. Equilibrium sedimentation and velocity experiments were performed on Beckman model XLA and XLI analytical ultracentrifuges equipped with absorbance or absorbance and interference optics, respectively. Absorbance as a function of radial distance was collected at 230 and 278 nm in a four cell rotor rated for 60 000 rpm. Interference as a function of radial distance was collected with a digital camera in an eight cell rotor rated for 50 000 rpm. Samples were prepared as described for SEC. The final buffer contained 10 mM potassium phosphate, 0.2 mM K_2-EDTA, and 1 mM DTE at pH 7.80. Integer values of urea concentrations ranging 0−8 M were examined at 25 °C. All samples were allowed to equilibrate at 25 °C for 1 h and the final protein concentration was 10−20 μM. Equilibrium sedimentation experiments were run at 32 000 rpm for ∼36 h until equilibrium was reached. Velocity sedimentation experiments were run at the maximum speed for the rotor used with data collected every 10 min until the boundary had been completely depleted.

The equilibrium data were analyzed assuming ideal and nonideal conditions to confirm the molecular weight and the sedimentation coefficient, s_0, which was determined from the velocity data. The software Origin was used to analyze the equilibrium data. Velocity data were analyzed by the transport method (*29*) and the second moment/boundary spreading method (*30, 31*). Good agreement between s_0 determined by equilibrium and velocity experiments confirmed the accuracy of the measurements. The partial specific volume, υ, was calculated from theoretical values for the individual amino acids (*32*) under native and unfolded conditions and linearly extrapolated to intermediate urea concentrations. The calculated values for υ and s_0 together with literature values for solvent density, ρ, and viscosity, η_0, were used to determine the Stokes, R_s, radius of the protein using the following equation from Tanford (*33*):

$$R_s = \mathrm{MW}(1 - \upsilon\rho)/(6N\pi\eta_0 s_0) \quad (3)$$

where MW is the molecular weight and *N* is Avogadro's number.

Dynamic Light Scattering. Equilibrium dynamic light-scattering experiments were performed on a Biotage dp-801 single angle instrument equipped with a 675 nm laser source. The diffusion coefficient was determined by Nelder−Meade Simplex analysis of the autocorrelation function derived from the digital signal [Biotage dp-801 Users Manual, 1992 (Protein Solutions Inc., Charlottesville, VA)]. The diffusion coefficient was determined at three different protein concentrations, 1.0, 2.0, and 3.0 mg mL^{-1}, and linearly extrapolated to zero concentration to eliminate nonideal effects. The Stokes radius was then calculated using the Stokes−Einstein equation (*33*):

$$R_s = kT/6\pi\eta D^T \quad (4)$$

where *k* is the Boltzmann constant, *T* is the temperature, η is the solvent viscosity, and D^T is the translational diffusion coefficient. Samples were prepared in triplicate as described for SEC and equilibrated at 25 °C for 1 h. The buffer contained 10 mM potassium phosphate, 0.2 mM K_2EDTA, and 1 mM DTE at pH 7.80. The final protein concentrations ranged between 30 and 100 μM and three sets of data were collected for each sample.

Small-Angle X-ray Scattering. Equilibrium SAXS experiments were performed on beam line 15A, a small-angle installation where a stable stream of photons is produced by a bent-crystal and horizontally focusing monochromator equipped with a vertically focusing mirror, at the Photon Factory of the National Institute for High Energy Physics, Tsukuba, Japan (*34*). The signal was detected with a gas-filled linear position-sensitive proportional counter camera from Rigaku (Tokyo, Japan) having 512 channels of 0.368 mm width positioned 2.346 m from the sample.

The raw SAXS data were corrected for the contrast effect of urea using standard procedures (*35*). Prior to calculation of the radius of gyration, R_g, the left and right sides of the scattering intensity profiles were averaged and filtered using

singular value decomposition (SVD) to construct a least-squares approximated data set (*36*). Although only the first three SVD basis vectors were nonrandom, the first four vectors were used to reconstruct a filtered data set. This filtering method has the advantage of allowing one to visually inspect, via the remaining essentially random basis vectors, what is being removed from the data. The radius of gyration was calculated at each urea concentration with the following equation for the Guinier approximation:

$$I(Q) = I(0)e^{-(1/3)R_g^2Q^2} \quad (5)$$

where Q and $I(Q)$ are the momentum transfer and intensity at zero scattering angle, respectively (*35*). Q is defined as $Q = 4\pi \sin\theta/\lambda$, where 2θ is the scattering angle and λ is the wavelength of the X-rays, 1.503 Å. The region used for the Guinier analysis was 0.00285−0.0157 Å^{-1} between 0 and 2 M urea and 0.001−0.00367 Å^{-1} above 2 M urea. The R_g of the native state was used to estimate R_s utilizing a hydrated model for the molecular volume (*37*). For the unfolded form, R_s was estimated from an R_s/R_g ratio of 1.55 for typical random coils (*38*).

Samples were prepared as described for SEC. The final buffer contained 10 mM potassium phosphate and 0.2 mM K_2EDTA at pH 7.80. All samples were run at 25 °C after 1 h of equilibration. The final protein concentration was ∼200 μM and data were collected for 300 s at each urea concentration. Scattering profiles at 0, 3, and 8 M urea were also collected at several protein concentrations for extrapolations of the Guinier analysis to a protein concentration of zero.

Thermodynamic Analysis. The urea dependence of the R_s values from SEC was fit to both three- and four-state models previously described in detail by Stackhouse et al. (*39*) and Gualfetti et al. (*24*), respectively. For both of the thermodynamic models, the free energy of unfolding in the absence of denaturant was calculated assuming a linear dependence of the apparent free-energy difference on the denaturant concentration (*28*, *40*, *41*):

$$\Delta G^\circ_{XY} = \Delta G^\circ_{XY(H_2O)} + m_{XY}[\text{urea}] \quad (6)$$

where ΔG°_{XY} is the free-energy difference between species X and Y at a given urea concentration, $\Delta G^\circ_{XY(H_2O)}$ is the free-energy difference in the absence of denaturant, and m_{XY} reflects the sensitivity of ΔG°_{XY} to the urea concentration. The data were fit using Savuka version 5.1, an in-house nonlinear least squares program (*42*).

Thermodynamic analysis of the SAXS data was accomplished by applying SVD to the data matrix of unfiltered, contrast-corrected scattering profiles as a function of urea (*36*, *43*). The statistically significant urea-dependent basis vectors, determined by their singular values, autocorrelation, and randomness, were fit to a four-state equilibrium denaturation model:

$$\text{N} \leftrightarrows \text{I1} \; (\leftrightarrows (\text{I1})_2) \leftrightarrows \text{I2} \leftrightarrows \text{U} \quad (7)$$

assuming a linear dependence of the apparent free energy on denaturant concentration as above.

The dissociation constant for the I1 $\rightleftharpoons$ $(I1)_2$ equilibrium, $K_d = 1 \times 10^{-4}$ M^{-1}, was determined by UC at 3 M urea and assumed to be independent of denaturant concentration. The native and unfolded baselines were assumed to be independent of denaturant concentration and were fixed parameters. In an effort to minimize the number of adjustable parameters, the *Z*-value, a normalized measure of the similarity of an intermediate to the native ($Z = 0$) and unfolded forms ($Z = 1$) (*22*, *44*), of $(I1)_2$ and I1 were assumed to be identical. The scattering profile for I1 and $(I1)_2$, therefore, represents a composite of the two species.

The scattering profiles for the thermodynamic species were obtained by linear combination of the *Q*-dependent basis vectors according to the coefficients obtained from the thermodynamic analysis (*43*). Kratky plots of N, I1, I2, and U were obtained from these scattering profiles.

RESULTS

The apparent radius of αTS was determined at a series of urea concentrations by hydrodynamic methods, including size-exclusion chromatography (SEC), dynamic light scattering (DLS) and ultracentrifugation (UC), and small-angle X-ray scattering (SAXS). The hydrodynamic methods reflect the diffusional properties of hydrated proteins, i.e., the Stokes radius R_s, while SAXS reflect the geometric distribution of the atoms in the protein framework, i.e., the radius of gyration, R_g. Comparison of R_s and R_g highlights the role of solvent and denaturant in the dynamic behavior of a protein in solution.

Size-Exclusion Chromatography. The application of SEC to measuring the size of native and unfolded proteins requires a resin whose separation properties are not affected by the presence of high concentrations of chemical denaturants. Superose 12 satisfies this requirement (*13*) and is capable of resolving proteins whose molecular masses range 1−300 kDa.

A calibration curve was generated by measuring the elution volumes, V_{el}, of 12 proteins whose hydrodynamic radii in both their native and urea-denatured forms have been previously determined (*13*). A plot of migration rate ($1000/V_{el}$) vs Stokes radius (*45*) for each of these standards in their native forms in 0 M urea and their reduced, unfolded forms in 8 M urea is shown in Figure 2. The data from the native and unfolded conformations of the entire set of 12 standard proteins can be fit to a single linear equation:

$$1000/V_{el} = 0.749R_s + 53.8 \quad (8)$$

Thus, both well-folded, globular conformations and random coil forms of these proteins are resolved to a comparable extent on the Superose 12 column.

The elution profiles for αTS at a representative set of urea concentrations are shown in Figure 3. The elution volume decreased from 12.8 to 9.2 mL when the urea concentration was increased from 0 to 8 M urea. The implied progressive increase in the size of αTS reflects the denaturant-induced shift in the equilibrium unfolding reaction to favor partially or fully unfolded protein. The broadening of the profile near 3 M urea reflects the kinetics of the N $\rightleftharpoons$ I1 reaction (*17*), which is thought to be comparable to the exchange time for the column matrix. Intermediate exchange between stable conformers of different size has previously been observed to have similar effects on other proteins (*13*, *46*, *47*). To account for the effects of intermediate exchange, the elution volume for αTS near 3 M urea was calculated from the center

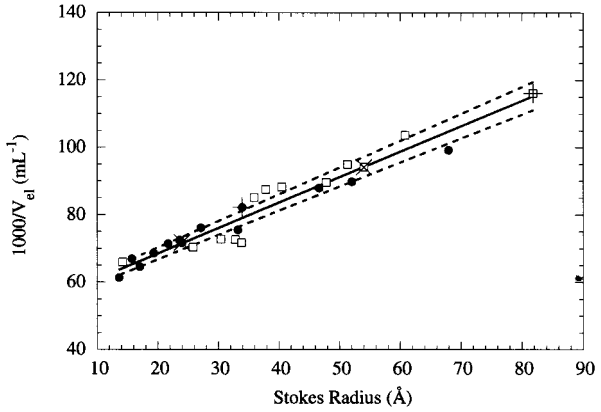

FIGURE 2: The migration rate vs Stokes radius for each of the twelve standard proteins used to calibrate the Superose 12 column. Each protein was run folded (●) in 0 M urea and unfolded (□) in 8 M urea. As an example, the two marked points (+) correspond to bovine serum albumin in the folded and unfolded form. Native and unfolded αTS (×) are also shown. The solid line represents a fit of all of the data to a linear equation, and the dashed lines represent the confidence intervals for 1 standard deviation. The buffer contained 100 mM potassium phosphate, 0.2 mM K_2EDTA and 1.0 mM DTE at pH 6.8. The concentrations of the proteins loaded on the column were all between 10 and 50 μM.

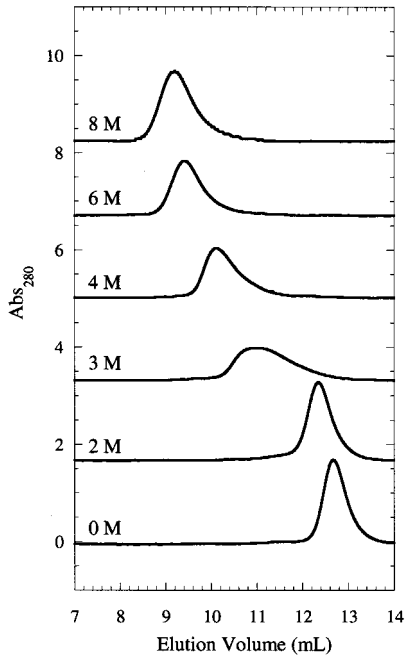

FIGURE 3: Representative elution profiles of αTS at various concentrations of urea. The data were collected at 25 °C in buffer containing 100 mM potassium phosphate, 0.2 mM K_2EDTA and 1.0 mM DTE at pH 7.8. The protein was loaded onto the Superose 12 column at a concentration of 10 μM.

of mass of the clearly non-Gaussian profile. The observation of near Gaussian profiles at all other urea concentrations permitted the V_{el} to be measured as the volume at the peak maximum.

The calibration curve shown in Figure 2 was then used to convert V_{el} for αTS to Stokes radius. The urea dependence of R_s for αTS is shown in Figure 4A. The R_s of the native state was found to be 24.2 Å (Table 1). The gradual, linear increase in R_s up to 2 M urea, with a slope of 0.35 Å M^{-1}, appears to reflect a response of the native conformation to increasing urea concentrations. This conclusion must be tempered by the magnitude of the estimated errors that are

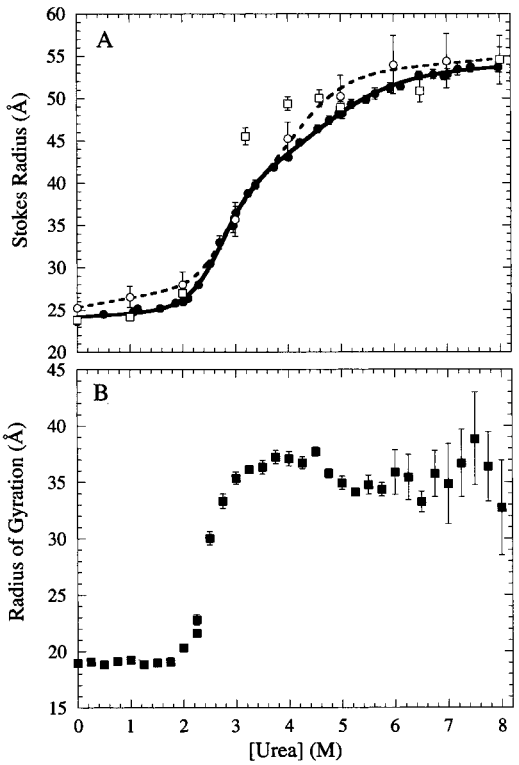

FIGURE 4: (A) The Stokes radius as a function of urea measured by SEC (●), UC (○), and DLS (□). The error bars for the SEC data represent the precision of the data and correspond to 1 standard deviation in the results of at least three trials at each urea concentration. The solid line represents the fit of the SEC data to a three-state model. The dashed line represents a fit of the UC data to a four-state model, fixing the free energies and *m* values to those obtained by optical techniques (Table 2) and fixing $Z = 1$ for the I2 ⇆ U transition. (B) Radius of gyration as a function of urea measured by SAXS (■). In addition to urea, the buffer also contained 10 mM potassium phosphate, 0.2 mM K_2EDTA and 1.0 mM DTE at pH 7.8 and 25 °C. For DLS, R_s was obtained by extrapolating the translational diffusion coefficient measured at 30, 60, and 90 μM to zero protein concentration (Materials and Methods). The protein concentration was 200 μM for SAXS.

determined by the accuracy of the calibration plot (Figure 2). Between 2 and 6 M urea, R_s undergoes a dramatic increase from ~27 to ~50 Å as αTS denatures. The apparent R_s at a given urea concentration is the weighted average of all molecular species present (*18*). The inflection near 3 M urea is indicative of the presence of a stable intermediate, designated as the I1 intermediate, that has been observed previously by optical methods and is part of the basis for the four-state equilibrium model (*22*, *24*). R_s then increases linearly to a value of 53.7 Å as the urea concentration is increased from 6 to 8 M where the unfolded form is highly populated (Table 1).

Ultracentrifugation. The Stokes radius of αTS was also measured by a combination of equilibrium and velocity sedimentation experiments. Ultracentrifugation offers an absolute measure of the size of a protein, avoiding the necessity of assuming that αTS does not have specific interactions with the Superose 12 resin and the requirement for a calibration curve.

Equilibrium sedimentation experiments were performed at unit urea concentrations between 0 and 8 M urea; concentration profiles in 0, 3 and 5 M urea solutions are shown in Figure 5, panels A, C, and E. αTS was observed

Table 1: Apparent Radius of the Native, Intermediate, and Unfolded Forms of αTS

technique	(radius)	radius (Å) N[a]	I1[a]	I2[a]	U[a]
SEC	(R_s)	24.2 ± 0.2	37.9 ± 4.0	52.9 ± 1.7	53.7 ± 0.6
UC	(R_s)	25.2 ± 1.2	37.1 ± 2.7	53.0 ± 3.2	54.5 ± 2.9
DLS	(R_s)	23.8 ± 0.5	ND[d]	ND[d]	56.2 ± 3.3
SAXS	(R_g)	19.1 ± 1.4	ND[d]	ND[d]	34.4 ± 3.7
	(R_s)[b]	23.8 ± 1.8	ND[d]	ND[d]	53.3 ± 5.7
prediction	(R_s)	24.6[c]	ND[d]	ND[d]	47.8[d]

[a] The apparent radius for N, I1, I2, and U were deconvoluted from the data under conditions where each species was maximally populated, 0.0, 3.2, 5.0, and 8.0 M urea, respectively. For the SEC data, slopes of the native and unfolded baselines from the thermodynamic analysis were used to calculate the R_s of native at 3.2 M urea and unfolded at 5.0 M urea. The slopes for the UC data were obtained by fitting the R_s data as a function of urea with a four-state model, in which the thermodynamic parameters were fixed at the values obtained by optical techniques and shown in Table 2. All other techniques used the data at 0 and 8 M urea as the apparent sizes for native and unfolded under all conditions. [b] R_s and R_g values are correlated by 1.29:1 at 0 M urea and by 1.55:1 at 8 M urea based on calculations of typical rigid spheres and unfolded proteins, respectively (*38*). [c] Predictions from SEC data of standard proteins were performed using $\log(R_s) = -0.254 + 0.369 \times \log(\mathrm{MW})$ for N and $\log(R_s) = -0.657 + 0.524 \times \log(\mathrm{MW})$ for U (*13*). [d] ND = not determined.

to behave as a monomer with a molecular mass of 28.7 kDa in all but the 3 M urea solution. At 3 M urea, αTS is best described by a monomer−dimer equilibrium with a dissociation constant of 1.3×10^{-4} M^{-1}. Because the stable I1 intermediate is dominant at 3 M urea (*22*), the monomer−dimer equilibrium must reflect the self-association reaction of I1. At 10−20 μM, the concentration used for velocity sedimentation experiments, ∼7−15% of the population would exist as dimeric I1. Less than 5% of the I1 species would exist as dimer at the ∼3 μM protein concentration that elutes from the column in SEC.

The velocity sedimentation scans for the 0, 3, and 5 M urea samples are shown in Figure 5, panels B, D, and F. The boundary broadening at increasing urea concentration reflects the longer times required for centrifuging the sample in the more viscous high urea solutions. The 6 h segments of data shown provide a reference for the length of time required for each sample to sediment as a function of urea. This broadening increases to some extent the uncertainty in the measurement of the sedimentation coefficient.

The results of both equilibrium and velocity experiments were used to estimate the R_s for αTS, as described in the Materials and Methods. The urea dependence of R_s at unit urea concentrations is plotted in Figure 4A. The R_s determined for native αTS at 0 M urea by UC, 25.2 Å, is in very good agreement with that obtained from SEC, 24.2 Å (Table 1). As was observed by SEC, a small but consistent increase in R_s up to 2 M urea suggests that the apparent radius of the native state increases prior to the unfolding reaction. Between 2 and 6 M urea, R_s undergoes a substantial increase in size that parallels the increase observed by SEC. R_s is then constant between 6 and 8 M urea, within the estimated errors. The R_s at 8 M urea measured by UC, 54.5 Å, is also in good agreement with that by SEC, 53.7 Å. Although the R_s values determined by UC are consistently 1−2 Å larger than those determined by SEC, the close correspondence between the two measures of the Stokes radius means that the results of the calibrated technique, SEC, are accurate, within the limits imposed by the errors.

Dynamic Light Scattering. R_s of αTS as a function of the urea concentration was also measured by DLS (Figure 4A). The radius of the native state, 23.8 Å, agrees well with the values obtained by SEC and UC (Table 1). The significant increase in R_s between 2 and 6 M urea reflects the expansion that accompanies the unfolding of the protein through the I1 and I2 intermediates. The observation that R_s is significantly larger than the value obtained from SEC and UC between 3 and 4 M urea most likely reflects the self-association of the I1 intermediate; the dimeric form is predicted to comprise 25−50% of the population at the concentrations employed in these experiments. The absence of a further change in R_s above 6 M urea suggests that the sizes of I2 and U are similar within the rather large estimates of error for this technique. R_s of the unfolded form, 56.2 Å, agrees well with the values obtained by SEC and UC (Table 1).

Small-Angle X-ray Scattering. The radius of gyration, R_g, was measured by SAXS to determine the expansion of the polypeptide in the absence of a contribution from the solvent (Figure 4B). A representative set of Guinier plots for αTS is shown in Figure 6A. $I(0)$ is independent of urea, except in the region from 2 to 4 M where the I1 intermediate is populated (data not shown). The $\ln I(Q)$ vs Q^2 plots become distinctly nonlinear at larger scattering angles as the urea concentration is increased, decreasing the range of data that are useful for estimating R_g. As a result, the errors at high urea concentrations are significant (Figure 4B).

The radius of gyration for native αTS was found to be 19.1 Å (Table 1), in excellent agreement with an earlier SAXS measurement of 19.5 Å at 4 °C (*37*). The R_s value predicted from the hydration of a protein with this R_g value, 23.8 Å, also agrees well with R_s values from hydrodynamic techniques (Table 1). R_g appears to remain constant as the urea concentration is increased up to 2 M urea (Figure 4B). R_g then proceeds through a maximum value of ∼37 Å between 3 and 4 M urea where the dimeric form of the I1 intermediate is predicted to represent ∼50% of the population. The increase in the zero angle intensity, $I(0)$, observed in this same urea range provides further evidence for the aggregation of I1 (data not shown). Above 4 M urea, R_g decreases to a constant value of ∼35 Å between 5 and 8 M urea. Although the self-association property of the I1 intermediate near 3 M urea and the inherently large errors at high denaturant concentration decrease the reliability of the SAXS data above 2 M urea, the value of R_s estimated from R_g at 8 M urea, 53.3 Å, is remarkably similar to the values obtained from the more precise SEC and UC data (Table 1).

Expression of the scattering data in the form of a Kratky plot yields insights into the globularity of the protein chain as a function of urea (*35*). Plots of $Q^2 \times I(Q)$ as a function of Q at 0, 3, 5, and 8 M urea, where the N, I1, I2, and U species are significantly populated, respectively, are shown in Figure 6B. The 0 M profile has a maximum that is characteristic of a globular protein. At 3 M urea, the peak is less intense and shifted to lower Q, indicative of a less structured, compact species of larger size. A significant loss of globularity is evident in the 5 and 8 M data and suggest

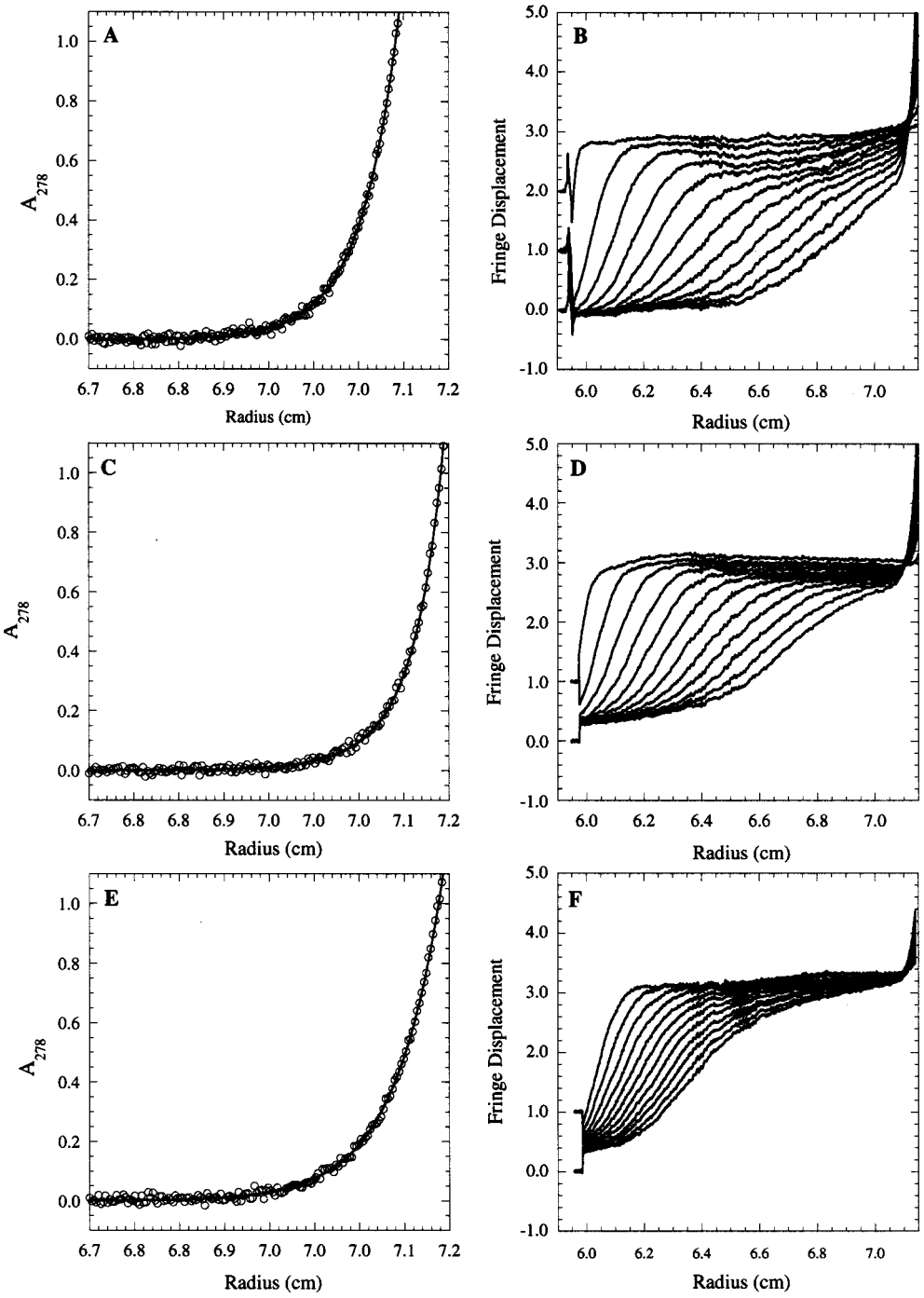

FIGURE 5: Representative equilibrium and velocity sedimentation data. (A, C, and E) Equilibrium sedimentation data collected for native αTS at 0, 3, and 5 M urea, respectively. For 0 and 5 M urea, the fits to a monomer with molecular mass 28.7 kDa are shown; for 3 M urea, the fit to a monomer−dimer equilibrium with a K_d of 1.3×10^{-4} M^{-1} is shown. (B, D, and F) Velocity sedimentation data collected at 0, 3, and 5 M urea, respectively. Scans taken every 30 min for the first 6 h are shown. For all velocity sedimentation experiments, only data where a detectable plateau existed before and after the boundary were fit. All samples were run at 25 °C in a buffer containing 10 mM potassium phosphate, 0.2 mM K_2EDTA and 1.0 mM DTE at pH 7.8. The protein concentration was 10−20 μM for all of the experiments.

that I2 and U are significantly unfolded at these urea concentrations. Although similar, a subtle difference between the 5 and 8 M urea data can be discerned at high scattering angles: the 5 M profile retains a maximum in the Kratky plot and slopes downward at high scattering angles whereas the 8 M profile appears to increase at high angles. The latter resembles the behavior expected for a more random coil-like structure (*43*). Because the 3 and 5 M urea data contain contributions from the N, I1, and I2 and the I1, I2, and U species, respectively, more accurate representations for each thermodynamic state can be obtained from an SVD analysis of the scattering data (see below).

Thermodynamic Analysis. Fits of the SEC- and SAXS-monitored unfolding reactions with equilibrium models can be used to test the possibility that these techniques can serve as an accurate monitor of the thermodynamic properties of the urea-induced unfolding reaction of αTS. A similar analysis was not performed on the UC or DLS data because the measurements of R_s were only taken at a limited number of urea concentrations. The apparent R_s measured by SEC

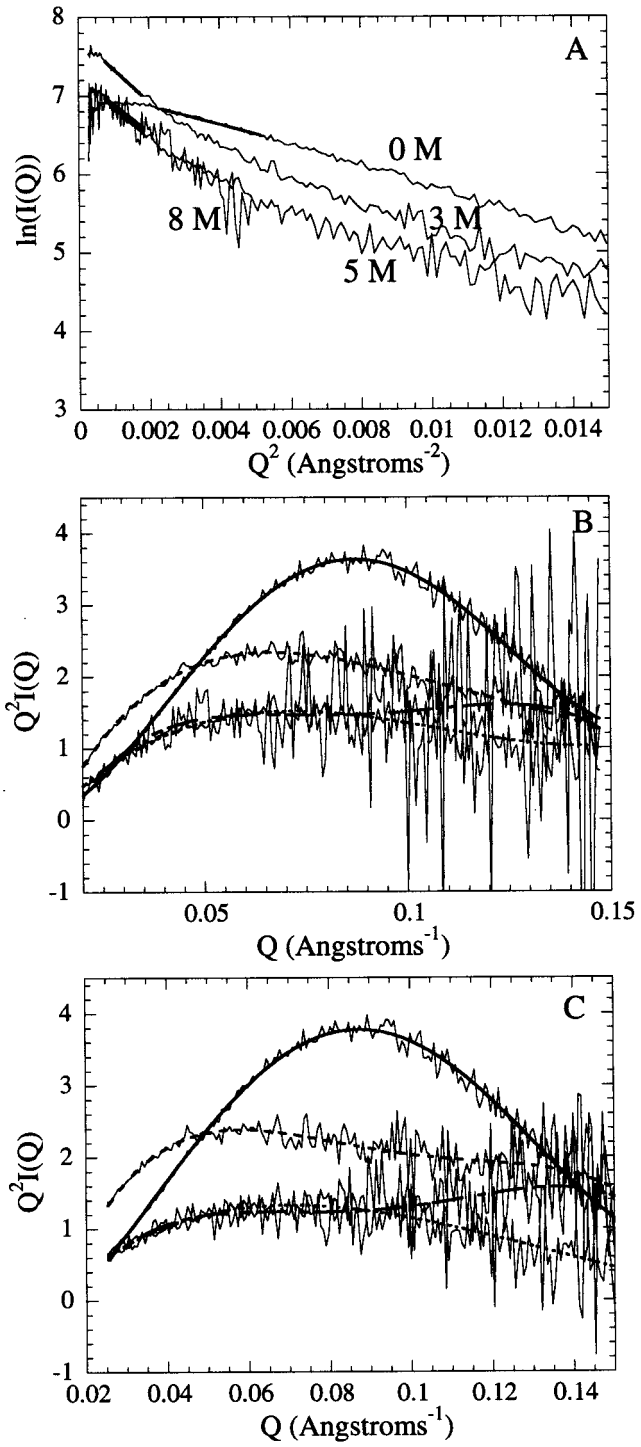

FIGURE 6: Small-angle X-ray scattering data collected as a function of urea at 200 μM protein. (A) Guinier plots of ln $I(Q)$ as a function of Q^2 at 0, 3, 5, and 8 M urea, where the linear portion used in the Guinier equation to calculate R_g is shown by the solid lines. (B) Kratky plots of $Q^2 \times I(Q)$ as a function of Q at 0 (solid line), 3 (dashed line), 5 (dotted line), and 8 (dashed−dotted line) M urea. (C) Kratky plots of the thermodynamic states, N (solid line), I1/$I1_2$ (dashed line), I2 (dotted line), and U (dashed−dotted line). The lines in panels B and C represent least squares fourth-order polynomials drawn to aid the eye. In addition to urea, the buffer contained 10 mM potassium phosphate, 0.2 mM K_2EDTA and 1.0 mM DTE at pH 7.8 and 25 °C.

as a function of urea (Figure 4A) was best fit by a model that involves three states, N ⇆ I ⇆ U. The free-energy changes in the absence of denaturant, $\Delta G^{\circ}_{(H_2O)}$, for the N ⇆ I and I ⇆ U transitions are 5.7 and 4.9 kcal mol^{-1}, respectively (Table 2). The sensitivities of these two transitions to denaturant, i.e., the m values, are −2.0 and −1.2 kcal mol^{-1} M^{-1}, respectively. The results are very similar to those observed for the N ⇆ I1 and I1 ⇆ I2 reactions in the four-state model determined by optical methods (Table 2). The observation that the I2 ⇆ U transition is adequately described by a linear dependence of the Stokes radius on the urea concentration implies that the apparent radii of these two states are indistinguishable by SEC. The SEC data, therefore, only reflect a fraction of the thermodynamic properties of αTS.

The validation of the Stokes radius as an accurate measure of the thermodynamic properties of αTS means that the observed R_s values represent a weighted average of the populated intermediates. Thus, the R_s for the I1 and I2 intermediates can be estimated in conjunction with the SEC data and thermodynamic results from previous optical experiments (*24*). A similar analysis was not performed with the SAXS data because of complications from the dimerization of I1 at high protein concentrations. From the free-energy changes, the relative populations of N, I1, I2, and U at 3.2 M urea are estimated to be 13, 73, 14, and 0%, respectively. When combined with the observed urea dependence of the R_s for the N state and the common urea dependence of the I2 and U states (Figure 4A), the observed R_s value for I1 at 3.2 M urea (Figure 4A) can be estimated (Table 1). Using an extrapolated value of 26.3 Å for the R_s of N at 3.2 M urea (using a slope of 0.35 Å M^{-1}), the Stokes radius of I1 is 37.9 Å, nearly 50% larger than that of the native state. The greater errors and the contribution of the self-associated forms of I1 make the estimates of R_s by DLS and SAXS unreliable. The R_s of the I2 state at 5.0 M urea, where the relative populations of the N, I1, I2, and U states are 0, 12, 67, and 21%, respectively, is estimated to be 52.9 Å.

In the case of the SAXS data, an SVD analysis of the raw scattering data yielded three significant components (Figure 7, panels A, C, and E). The corresponding urea dependences (Figure 7, panels B, D, and F) were fit to a four-state model that explicitly incorporates a dissociation constant of 1×10^{-4} M^{-1} for the dimeric form of the intermediate. In contrast to the Guinier analysis, which was not sensitive to the I2 ⇆ U transition, the SVD analysis incorporating the entire scattering profile showed statistically significant sensitivity to this transition. The large number of parameters required for a four-state fit, however, necessitated fixing the thermodynamic parameters for the I2 to U transition to those obtained from optical studies and allowing only a single additional parameter, the Z-value for I2, to remain adjustable in the least-squares optimization. The resulting thermodynamic parameters, shown in Table 2, were used along with the SVD basis vectors to reconstruct the Kratky plots for the thermodynamic species, N, I1/$(I1)_2$, I2, and U (Figure 6C). As inferred from the Kratky plots at 0, 3, 5, and 8 M urea, the plots of N, I1, and I2 indicate that these species are globular while U is more random coil-like.

DISCUSSION

Native State of αTS. The native state has a Stokes radius of ∼24 Å, as measured by SEC, UC, and DLS, and a radius of gyration measured by SAXS of 19.1 Å. The excellent

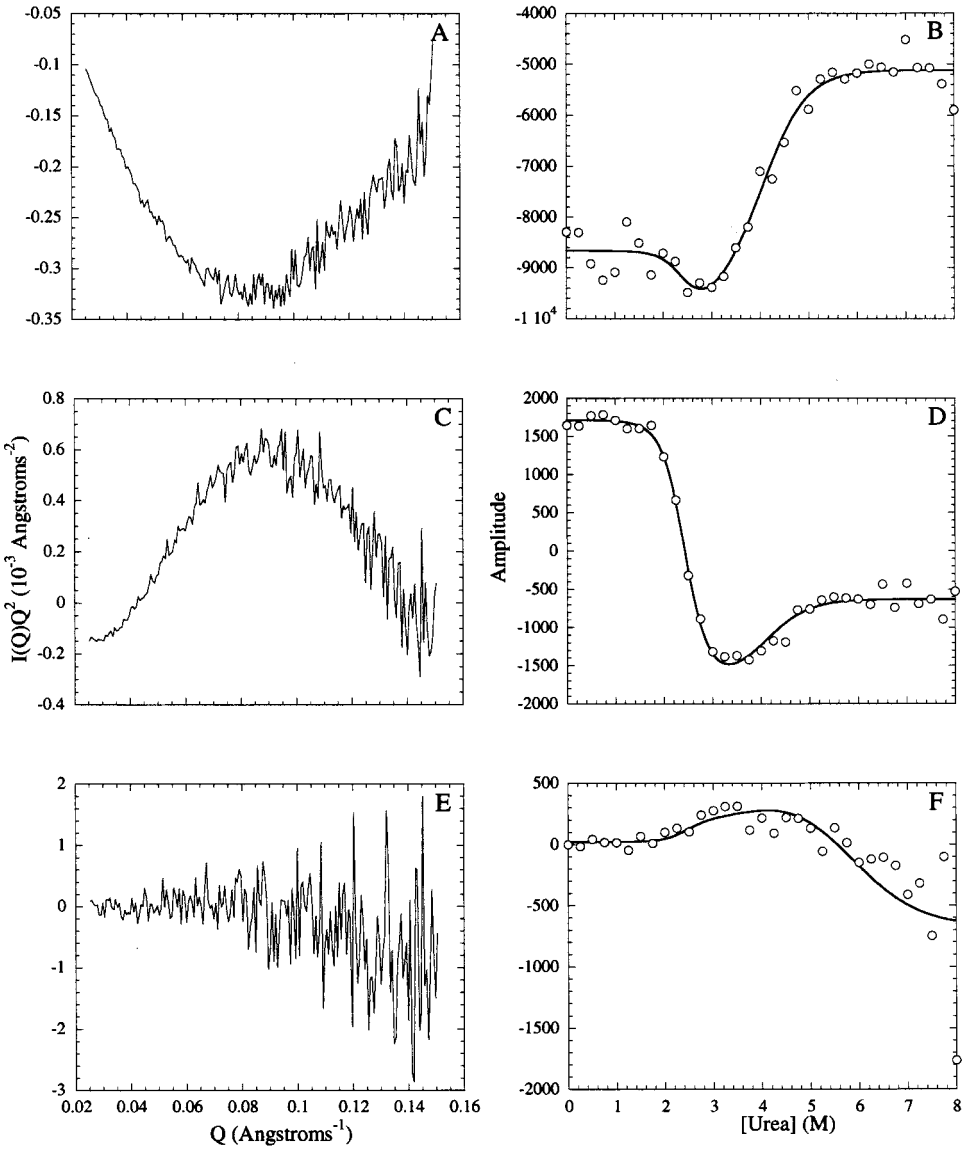

FIGURE 7: Thermodynamic analysis of SAXS data. (Left panels) Kratky plots of the first (A), second (C), and third (E) most significant u-vectors, with relative weights of 75.9, 12.9, and 4.4%, respectively. (Right panels) Urea dependence of the corresponding first (B), second (D), and third (F) most significant SVD v-vectors. The data were fit to a four-state equilibrium model that accounted for the dimerization of I1 with a K_d of 1×10^{-4} M^{-1} to yield the thermodynamic parameters shown in Table 2.

Table 2: Thermodynamic Parameters for the Unfolding of αTS

technique	$\Delta G^\circ_{N/I1(H_2O)}$ (kcal mol^{-1})	$-m_{N/I1}$ (kcal mol^{-1} M^{-1})	$\Delta G^\circ_{I1/I2(H_2O)}$ (kcal mol^{-1})	$-m_{I1/I2}$ (kcal mol^{-1} M^{-1})	$\Delta G^\circ_{I2/U(H_2O)}$ (kcal mol^{-1})	$-m_{I2/U}$ (kcal mol^{-1} M^{-1})
optical	6.0	2.2	4.5	1.1	4.9	0.84
(global fit)[a]	(0.1)[b]	(0.04)	(0.2)	(0.04)	(0.4)	(0.06)
SEC	5.7	2.0	4.9	1.2	ND[c]	ND[c]
	(0.4)	(0.1)	(0.5)	(0.2)		
SAXS	5.3	2.0	3.1	0.9	5.0[d]	0.85[d]
	(1.0)	(0.4)	(0.9)	(0.2)		

[a] Simultaneous fit of data from fluorescence total intensity, fluorescence anisotropy, and circular dichroism spectroscopy (*24*). Individual data sets were weighted by the RMS deviation within the native baseline. [b] Errors in parentheses represent one standard deviation from an uncorrelated error analysis. [c] ND = not determined. [d] Parameters for the I2 to U transition were fixed to values obtained from optical studies of the unfolding reaction (*24*).

agreement with the values predicted from the X-ray structure of αTS in the tetrameric $\alpha_2\beta_2$ tryptophan synthase complex suggests that isolated αTS maintains a very similar structure. The small, linear increase in R_s for the native conformation observed by SEC, UC, and, possibly, DLS between 0 and 2 M urea could reflect either a slight expansion of the folded structure in low concentrations of denaturant or the effect of urea binding to the surface of the protein (*5*, *48*). The former explanation is less likely because the far-UV CD and fluorescence spectra indicate that the secondary and tertiary structures are not perturbed in this range (*49*). It is more likely that urea bound to the surface of the native protein results in an apparently larger R_s by limiting the accessibility to pores in the Superose 12 resin or by slowing translational

diffusion in the sedimentation and DLS experiments. The observation that low concentrations of urea also increase the number of well-ordered water molecules at the surface of dihydrofolate reductase in its crystalline form suggests that structured water may also play a role in the apparent increase in size of native αTS (*48*).

I1 State. As αTS is induced to unfold by urea, the I1 intermediate becomes highly populated in 3.2 M urea (*24*). Other partially folded intermediates that have been detected by SEC include molten globule forms of bovine carbonic anhydrase β and α-lactalbumin (*13*). The R_s of both partially folded forms are ∼10% larger than the native state, typical of molten globule forms of other proteins (*50, 51*); the increases account for ∼25% of the change in size accompanying unfolding. On the basis of the SEC data, it appears that I1 is about 50 ± 24% larger than the native form if both are compared at 3.2 M urea. If the UC data are considered, I1 is about 28 ± 10% larger than N under the same conditions. The large apparent discrepancy between the SEC and UC techniques reflects the differing estimates for the native form and its dependence on urea, which affects the absolute magnitude of the R_s for I1. By either measure, this significant increase in size is accompanied by the selective disruption of the exterior helical segments and the exposure of buried tyrosines to solvent (*24*). The retention of tertiary packing near at least some of the buried aromatic residues, demonstrated by both near-UV CD and steady-state tyrosine FL anisotropy (*24*), however, suggests that I1 does not satisfy the classic definition of a molten globule (*8*). An alternative explanation is that some portions of the protein are disordered while other portions retain a native-like packing environment. This conjecture is consistent with a model where the surface helices are preferentially unfolded or frayed while the core β-strands are largely intact and stabilized by side-chain−side-chain interactions. Mutational analysis suggests that the side chains in strands 1 and 8 are not yet in contact, implying that I1 may have an open, sheetlike form (*52*).

A previous calorimetric and spectroscopic study of the I1 intermediate (*53*) concluded that I1 is a molten globule state, lacking a measurable endotherm and containing substantial secondary structure but little discernible tertiary structure. As acknowledged by those investigators, the interpretation of calorimetric data in the presence of urea is complicated by the exothermic binding of the denaturant to partially or fully unfolded proteins. Furthermore, the analysis of the calorimetric and NMR data is complicated by the demonstrated tendency of this intermediate to self-associate significantly at the high concentrations required for these methods. Finally, the observation of a near-UV CD signal indicative of nonrandom side chain packing in the I1 species by these investigators (*53*) suggests that I1 more closely resembles a structured molten globule (*54*) that retains residual tertiary structure.

The identification of the I1 intermediate from αTS as an on-pathway kinetic intermediate (*17*) provides the opportunity to compare the sizes of the stable and transient forms. Time-resolved fluorescence anisotropy studies of 1,8-anilinonaphthalene sulfonate bound to αTS during its folding reaction (*5*) have shown the R_s of N and I1 in 0.6 M urea to be 25.9 and 30.1 Å, respectively. Although the estimated size of the native conformation is in excellent agreement with those reported in the present communication, there is at least an apparent discrepancy between the R_s values for the stable and transient forms of I1 (Table 1). This discrepancy may have several possible explanations. (1) The R_s for the transient form of I1 was measured by TRA in 0.6 M urea while the equilibrium measurements were performed in 3.2 M urea. Equilibrium measurements of the size of I1 by TRA in ∼3 M urea resulted in a R_s value of ∼32 Å (*5*), larger than the 30.1 Å for the transient form of I1 but still smaller than the SEC value at 3.2 M urea, 37.9 ± 4.0 Å. (2) Mechanistic studies of the folding reaction (*17*) show that I1 is comprised of four kinetic species that interconvert through isomerization or rearrangement reactions. Differential binding of ANS to these forms or different quantum yields when bound could serve to emphasize the rotational dynamics of a smaller member of this set. SEC and UC monitor the population-weighted average of the entire set whose average radius might be larger. (3) The shape of I1 may affect all of these measurements; however, TRA data suggest that the molecule is roughly spherical (O. Bilsel, J. M. Beechem, and C. R. Matthews, unpublished results). (*4*) The partially folded nature of the intermediate may lead to a lower estimated value of R_s by TRA if the region of I1 to which ANS binds is capable of rotating more rapidly than the global motion of the entire polypeptide. (*5*) The calibration curve for the SEC experiment may misrepresent R_s of partially folded forms because it only utilizes native and unfolded proteins. However, the good agreement between the SEC and UC measurements of R_s, under conditions where I1 is highly populated, make this explanation seems unlikely. Regardless of the discrepancies in the apparent radius of I1, both equilibrium and kinetic techniques reveal that the radius of I1 is significantly smaller than I2 or U and larger than N.

I2 State. The I2 state at 5.0 M urea has the same R_s as the unfolded form in this solvent, although both are ∼3 Å smaller than the U state at 8 M urea. Although the errors in R_g from the SAXS experiment preclude a judgment of its dependence on urea, the Kratky analysis of the scattering profiles suggests that I2 has a more globular shape than U. The small contraction in R_s and the development of long-range order is accompanied by the formation of a global hydrophobic cluster that is responsible for the distinctive thermodynamic properties of the I2 state (*25, 55*). Specifically, the enthalpy and entropy changes for the I2 → U reaction are large and negative: $\Delta H° = -44$ kcal mol^{-1} and $\Delta S° = -174$ cal mol^{-1} K^{-1} at 25 °C. Contiguous strings of nonpolar side chains, such as those often found in β-strands, might be sufficiently insoluble even in highly denaturing solutions that they would spontaneously form hydrophobic clusters in the absence of intrachain hydrogen bonding, i.e., without the simultaneous formation of a β-sheet. The equivalent Stokes radius for I2 and U implies that the reorganization required to form this cluster occurs without a physical contraction of the conformational ensemble, beyond that experienced by the U state. Presumably, the loops, turns and helices that link the strands would remain unstructured on the exterior of this cluster and prevent it from forming large aggregates with other molecules.

Other examples of residual structure in highly denaturing solutions include the 434 repressor. Four near-neighbor residues in the sequence form a hydrophobic cluster that persists in 7 M urea without the hydrogen-bonding pattern

found in the native, helical structure (*56*). Also, a highly unfolded intermediate has been detected in cytochrome *c* by SAXS. Two forms with indistinguishable R_g values, U1 and U2, show different scattering profiles at high angle. The longer range coherent scattering observed for U1 implies that it has a globular shape while U2 has a random coil-like shape. Significant stability and an appreciable *m* value suggest that the residual structure in the U1 species for cytochrome *c* involves the burial of substantial surface area relative to U2 (*43*). Although the all-helix fold of cytochrome *c* contains interspersed hydrophilic and hydrophobic side chains, the covalently bound heme group may provide a nucleation site for the formation of a hydrophobic cluster.

U State. The unfolded state of αTS has a Stokes radius of ~55 Å in 8 M urea by SEC, UC, and DLS. The Stokes radius of U estimated from R_g at 8 M urea from SAXS measurements and based on typical unfolded proteins, 53.3 Å (Table 1), is in excellent agreement with the hydrodynamic measurements. Although these values are slightly larger than expected for the random coil form of a protein containing 268 amino acids (*13*), 47.8 Å, the urea dependence for the I2 and U states detected by SEC could account for the apparent discrepancy. Linear extrapolation of the unfolded baseline region above 6 M urea to the absence of urea yields a Stokes radius of 50.5 Å, in better agreement with the predicted value. Polynomial extrapolations of the sort recently proposed for staphylococcal nuclease (*18*) would lead to an even smaller estimated value for R_s, perhaps closer to the predicted value.

The view of unfolded states as compact, denatured ensembles of highly heterogeneous polymers (*18, 19*) suggests that even under strongly unfolding conditions, most if not all proteins may not behave as ideal random coils. This supposition is supported for αTS by the presence of residual structure in high concentrations of denaturant detected by optical methods, such as near-UV CD spectroscopy (*53, 57*) and fluorescence (*24*).

Stokes Radius as a Measure of the Size of Conformational Ensembles for Stable Intermediates in the Folding of αTS. Using a combination of hydrodynamic and scattering techniques to determine R_s, it has been shown that the size of the conformational ensembles representing the native, intermediate, and unfolded forms decrease in parallel with the formation of secondary and tertiary structure. Specifically, the formation of a hydrophobic cluster that stabilizes I2 in the near absence of secondary and tertiary structure occurs only with the small contraction also experienced by the unfolded form (Table 1). Apparently, the amphipathic segments that link the hydrophobic sequences thought to be responsible for the cluster (*23, 25, 55*) protrude into the solvent and maintain a Stokes radius that is very similar to the unfolded form. Thus, R_s can only be taken to be an approximate measure of the size of a conformational ensemble. It is interesting to note that the formation of a hydrophobic cluster without the simultaneous development of higher order structure would minimize the conformational entropy penalty; the observed increase in entropy accompanying the conversion of U to I2 (*25*) must reflect a dominant role for solvent release from nonpolar side chains in the cluster.

By contrast, the formation of I1 and N is accompanied by substantial decreases in R_s and similar, large fractional changes in tyrosine fluorescence intensity, tyrosine absorbance, and the far-UV CD signal at 222 nm (*24*). The progressive development of secondary and tertiary structure in these states occurs along with a substantial contraction of their conformational ensembles.

ACKNOWLEDGMENT

We thank Lubomir Kovac for his technical assistance with the Beckman XL-A at the University of Texas Medical Branch at Galveston. We thank Dr. Gregory Farber for the use of his light scattering instrument and Dr. David Lambright for providing an early version of Savuka. We also thank an anonymous reviewer for helpful comments on SEC.

REFERENCES

1. Shortle, D., Chan, H. S., and Dill, K. A. (1992) *Protein Sci. 1*, 201−215.
2. Karplus, M., and Weaver, D. L. (1976) *Nature 260*, 404−406.
3. Smith, L. J., Fiebig, K. M., Schwalbe, H., and Dobson, C. M. (1996) *Folding Des. 1*, R95−R106.
4. Cantor, C. R., and Schimmel, P. R. (1980) Techniques for the Study of Biological Structure and Function. *Biophysical Chemistry*, Vol. II, W. H. Freeman Co., New York.
5. Bilsel, O., Yang, L., Zitzewitz, J. A., Beechem, J. M., and Matthews, C. R. (1999) *Biochemistry 38*, 4177−4187.
6. Lillo, M. P., Beechem, J. M., Szpikowska, B. K., Sherman, M. A., and Mas, M. T. (1997) *Biochemistry 36*, 11261−11272.
7. Lillo, M. P., Szpikowska, B. K., Mas, M. T., Sutin, J. D., and Beechem, J. M. (1997) *Biochemistry 36*, 11273−11281.
8. Kuwajima, K. (1989) *Proteins 6*, 87−103.
9. Ptitsyn, O. B. (1995) *Adv. Protein Chem. 47*, 83−229.
10. Goto, Y., Takahashi, N., and Fink, A. L. (1990) *Biochemistry 29*, 3480−3488.
11. Goto, Y., and Fink, A. L. (1994) *Methods Enzymol. 232*, 3−15.
12. Kataoka, M., Nishii, I., Fujisawa, T., Ueki, T., Tokunaga, F., and Goto, Y. (1995) *J. Mol. Biol. 249*, 215−228.
13. Uversky, V. N. (1993) *Biochemistry 32*, 13288−13298.
14. Jennings, P. A., and Wright, P. E. (1993) *Science 262*, 892−896.
15. Raschke, T. M., and Marqusee, S. (1997) *Nat. Struct. Biol. 4*, 298−304.
16. Chan, H. S., and Dill, K. A. (1998) *Proteins 30*, 2−33.
17. Bilsel, O., Zitzewitz, J. A., Bowers, K. E., and Matthews, C. R. (1999) *Biochemistry 38*, 1018−1029.
18. Baskakov, I. V., and Bolen, D. W. (1998) *Biochemistry 37*, 18010−18017.
19. Dill, K. A., and Shortle, D. (1991) *Annu. Rev. Biochem. 60*, 795−825.
20. Hyde, C. C., Ahmed, S. A., Padlan, E. A., Miles, E. W., and Davies, D. R. (1988) *J. Biol. Chem. 263*, 17857−17871.
21. Nichols, B. P., and Yanofsky, C. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 5244−5248.
22. Matthews, C. R., and Crisanti, M. M. (1981) *Biochemistry 20*, 784−792.
23. Saab-Rincón, G., Froebe, C. L., and Matthews, C. R. (1993) *Biochemistry 32*, 13981−13990.
24. Gualfetti, P. J., Bilsel, O., and Matthews, C. R. (1999) *Protein Sci. 8*, 1623−1635.
25. Saab-Rincón, G., Gualfetti, P. J., and Matthews, C. R. (1996) *Biochemistry 35*, 1988−1994.
26. Kirschner, K., Wiskocil, R., Foehn, M., and Rezeau, L. (1975) *Eur. J. Biochem. 60*, 513−523.
27. Miles, E. W. (1970) *J. Biol. Chem. 245*, 6016−6025.
28. Pace, C. (1986) *Methods Enzymol. 131*, 266−280.
29. Schachman, H. K. (1959) *Ultracentrifugation in Biochemistry*, Academic Press, New York.
30. Muramatsu, N., and Minton, A. P. (1988) *Anal. Biochem. 168*, 345−351.
31. Goldberg, R. J. (1953) *J. Phys. Chem. 57*, 194.

32. Lee, J. C., and Timasheff, S. N. (1974) *Arch. Biochem. Biophys. 165*, 268−273.
33. Tanford, C. (1961) *Physical Chemistry of Macromolecules*, John Wiley & Sons, New York.
34. Amemiya, Y., Wakabayashi, K., Hamanaka, T., Wakabayashi, T., Matsushima, T., and Hashizume, H. (1983) *Nuclear Instr. Methods 208*, 471−477.
35. Glatter, O., and Kratky, O. (1982) *Small Angle X-ray Scattering*, Academic Press, London.
36. Henry, E. R., and Hofrichter, J. (1992) *Methods Enzymol. 210*, 129−192.
37. Wilhelm, P., Pilz, I., Lane, A. N., and Kirschner, K. (1982) *Eur. J. Biochem. 129*, 51−56.
38. Damaschun, G., Damaschun, H., Gast, K., Gernat, C., and D., Z. (1991) *Biochem. Biophys. Acta 1078*, 289−295.
39. Stackhouse, T., Onuffer, J., Matthews, C., Syed, A., and Miles, E. (1988) *Biochemistry 27*, 824−832.
40. Matthews, C. R. (1987) *Methods Enzymol. 154*, 498−511.
41. Schellman, J. A. (1978) *Biopolymers 17*, 1305−1322.
42. Zitzewitz, J. A., Bilsel, O., Luo, J., Jones, B. E., and Matthews, C. R. (1995) *Biochemistry 34*, 12812−12819.
43. Segel, D. J., Fink, A. L., Hodgson, K. O., and Doniach, S. (1998) *Biochemistry 37*, 12443−12451.
44. Tanford, C. (1968) *Adv. Protein Chem. 23*, 121−282.
45. Davis, L. C. (1983) *Chromatogr. Sci. 21*, 214.
46. Shalongo, W., Jagannadham, M., and Stellwagen, E. (1993) *Biopolymers 33*, 135−145.
47. Shalongo, W., Heid, P., and Stellwagen, E. (1993) *Biopolymers 33*, 127−134.
48. Dunbar, J., Yennawar, H. P., Banerjee, S., Luo, J., and Farber, G. K. (1997) *Protein Sci. 6*, 1727−1733.
49. Beasty, A. M., Hurle, M. R., Manz, J. T., Stackhouse, T., Onuffer, J. J., and Matthews, C. R. (1986) *Biochemistry 25*, 2965−2974.
50. Zhang, Y., and Gray, R. D. (1996) *J. Biol. Chem. 271*, 8015−8021.
51. Gast, K., Damaschun, H., Misselwitz, R., Muller-Frohne, M., Zirwer, D., and Damaschun, G. (1994) *Eur. Biophys. J. 23*, 297−305.
52. Tsuji, T., Chrunyk, B. A., Chen, X., and Matthews, C. R. (1993) *Biochemistry 32*, 5566−5575.
53. Ogasahara, K., Matsushita, E., and Yutani, K. (1993) *J. Mol. Biol. 234*, 1197−1206.
54. Baldwin, R. L. (1991) *Chemtracts-Biochem. Mol. Biol. 2*, 379−389.
55. Zitzewitz, J. A., Gualfetti, P. J., Perkons, I. A., Wasta, S. A., and Matthews, C. R. (1999) *Protein Sci. 8*, 1200−1209.
56. Neri, D., Billeter, M., Wider, G., and Wuthrich, K. (1992) *Science 257*, 1559−1563.
57. Zitzewitz, J. A., and Matthews, C. R. (1999) *Biochemistry 38*, 10205−10214.

BI991296S