Published in final edited form as:

J Mol Biol. 2008 March 14; 377(1): 294–306.

NMR Analysis of Partially-Folded States and Persistent Structure in the Alpha Subunit of Tryptophan Synthase:

Implications for the Equilibrium Folding Mechanism of a 29 kDa TIM Barrel Protein

Ramakrishna Vadrevu, Ying Wu, and C. Robert Matthews*

Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, Worcester, MA 01605

Abstract

Structural insights into the equilibrium folding mechanism of the alpha subunit of tryptophan synthase (α TS) from *E. coli*, a $(\beta\alpha)_8$ TIM barrel protein, were obtained with a pair of complementary NMR spectroscopic techniques. The secondary structures of rare high-energy partially-folded states were probed by native-state hydrogen exchange NMR analysis of main chain amide hydrogens. 2D HSQC NMR analysis of several ^{15}N -labeled nonpolar amino acids was used to probe the side chains involved in stabilizing a highly-denatured intermediate that is devoid of secondary structure. The dynamic broadening of a subset of isoleucine and leucine side chains and the absence of protection against exchange showed that the highest energy folded state on the free energy landscape is stabilized by a hydrophobic cluster lacking stable secondary structure. The core of this cluster, centered near the N-terminus of α TS, serves as a nucleus for the stabilization of what appears to be non-native secondary structure in a marginally-stable intermediate. The progressive decrease in protection against exchange from this nucleus towards both termini and from the N-termini to the C-termini of several β -strands is best described by an ensemble of weakly-coupled conformers. Comparison with previous data strongly suggests that this ensemble corresponds to a marginally-stable off-pathway intermediate that arises in the first few milliseconds of folding and persists under equilibrium conditions. A second, more stable, intermediate, which has an intact β -barrel and a frayed α -helical shell, co-exists with this marginally-stable species. The conversion of the more stable intermediate to the native state of α TS entails the formation of a stable helical shell and completes the acquisition of the tertiary structure.

Keywords

Hydrogen exchange; folding intermediates; hydrophobic cluster; HSQC; secondary structure

The equilibrium unfolding reactions of globular proteins, induced by a variety of perturbants such as chemical denaturants or extremes of temperature and pH, are usually highly cooperative processes. Often only the native, biologically-active conformation and the denatured state are measurably populated as the system progressively shifts between these well-defined thermodynamic states under the application of the perturbation. If the unfolding reaction is fully reversible and well-described by a two-state model, it is possible to extract fundamental

* Author to whom correspondence should be addressed. Phone: (508) 856 2251. Fax: (508) 856 8358. Email: c.robert.matthews@umassmed.edu

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

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

parameters such as the changes in free energy, enthalpy, entropy and heat capacity.¹ Although the simplicity of folding reactions at equilibrium has been a boon for the measurement of thermodynamic properties, the absence of partially-folded states has been a major impediment towards the understanding of the molecular events that dictate the rapid and efficient conversion of a highly-disordered denatured state to a tightly-packed native conformation for many proteins.

Kinetic studies of folding reactions often reveal the presence of transient intermediates under strongly folding conditions for proteins containing more than 100 amino acids.²⁻⁵ Unfortunately, the short lifetimes, marginal stabilities and dynamic properties of these intermediates, which might be expected to reveal the relationships between sequence, structure and folding, make it very challenging to obtain the desired high-resolution structural information during folding reactions. The common observation that the rate-limiting step in most folding reactions involves the formation of the native state means that preceding reactions, i.e., those related to the formation and disruption of intermediates, are likely to be at quasi-thermodynamic equilibrium under strongly folding conditions. Thus, efforts to determine the structures of intermediates under equilibrium conditions favoring the native state have the potential to provide key insights into the molecular mechanisms of folding reactions.⁶⁻⁹

Direct detection of these marginally-stable intermediates, however, is difficult because they only occupy a very small fraction of the population in the absence of denaturant or at room temperature and neutral pH. Fortunately, the protection of main chain amide hydrogens against exchange with solvent deuterium by hydrogen bonding with carbonyl oxygens provides an indirect readout of persistent hydrogen bonding patterns in rare, high energy folding intermediates.^{10,11} When combined with assignments of amide hydrogen cross peaks in a Heteronuclear Single Quantum Coherence (HSQC) NMR spectrum, it has been possible to assess the residual secondary structure of rare partially-folded states at the level of individual amino acids.⁸⁻¹⁰

The alpha subunit of tryptophan synthase (α TS) from *E. coli* is an excellent vehicle for exploring the application of the native-state HX-NMR technique to structural analysis of the intermediates in a complex folding reaction. α TS is a member of the TIM barrel super-family, one of the most common motifs in biology.^{12,13} The 8 sequential parallel β -strands form a cylindrical structure that is surrounded and solubilized by a shell of 8 α -helices that alternate in sequence with the strands.¹⁴

The combination of a battery of spectroscopic methods under equilibrium conditions¹⁵⁻¹⁷ and a global analysis of the kinetic folding mechanism¹⁸ has produced a detail map of the folding free energy surface for α TS. Prominent on this surface are the existence of three distinct intermediates, I1, I2 and I_{BP}, that are distinguished by their structural and thermodynamic properties. Their linkages with each other and with the native, N, and unfolded, U, states are shown in Scheme 1.

A reaction coordinate diagram summarizing their free energies relative to the native state at pH 7.8 and 25 °C in the absence of denaturant is shown Figure 1. The I_{BP} species is drawn to the right of the U species in Figure 1 because, in the kinetic folding mechanism, I_{BP} serves as an off-pathway kinetic trap.¹⁸ I_{BP} must back-track through the less well-folded I2 and/or U species before α TS can fold productively through the I1 intermediate to the N state. Although I_{BP} is an off-pathway kinetic species, it is populated, ~10%, at equilibrium under moderately denaturing conditions, 3 M urea. I1, by contrast, forms ~75% of the population at equilibrium in 3 M urea¹⁸, and the I2 intermediate comprises 70% of the population at 5 M urea. The three proline isomerization reactions that limit folding under conditions favoring the native conformations¹⁸ are only apparent in the kinetic folding mechanism. The pairs of cis/trans

isomers are known to modulate, in minor ways, the free energies of each of the principal states¹⁸ in the reaction coordinate diagram.

The I1 intermediate retains ~50% of the far-UV CD signal at 222 nm; SVD analysis of the CD data suggests that the β -barrel is intact and that the α -helices are less well structured.¹⁷ The I_{BP} intermediate also retains about 50% of the far-UV CD signal of the native state. However, I_{BP} is less stable than its I1 counterpart and is less responsive to denaturation by urea. The decreased sensitivity of the unfolding transition to the urea concentration, i.e., the m-value, implies that I_{BP} is either a less well-packed structure than I1¹⁸ or that additional partially-folded states appear as I_{BP} is induced to unfold at higher urea concentrations. The I2 intermediate is stabilized by the hydrophobic effect,^{16,19} and its ellipticity under strongly-denaturing conditions, 5 M urea, is indistinguishable from that of the U state at 8 M urea. The observation of substantial far-UV CD signals for both the I1 and I_{BP} species makes them potential candidates for protection against exchange with solvent for their amide hydrogens. The absence of ellipticity in the I2 species makes it a potential vehicle for the exchange of the most resistant amide hydrogens.

The availability of assignments for a majority of cross peaks in the ¹H-¹⁵N HSQC NMR spectrum for α TS²⁰ provided an opportunity to explore the protection against exchange for its rare high-energy states via the native-state HX experiment at low denaturant concentrations. A complementary ¹H-¹⁵N HSQC analysis of α TS containing selectively ¹⁵N-labeled isoleucine, leucine, phenylalanine or valine provided insights into the composition of the core of the hydrophobic cluster defining the I2 state under highly-denaturing conditions. Comparison of these results with the properties of the intermediates known to reside on the folding free energy surface provides useful insights into the development of structure and stability during the equilibrium folding reaction of α TS.

Results

Native state HX NMR experiment

The previous assignment of ~85% of the 249 backbone amide hydrogen HSQC cross peaks in α TS²⁰ (268 amino acids less 19 prolines) provided the starting point for the native-state hydrogen exchange experiment (Figure 2A). These assignments include 85% of the 130 residues in the α -helices and 92% of the 37 residues in the β -strands, likely positions for protection against exchange in partially-folded forms of α TS.

Exchange of the amide hydrogens for deuterium was measured by recording the 2D ¹H-¹⁵N HSQC spectra as a function of time after dissolving lyophilized protein into ²H₂O at pD 7.8, 25 °C at a series of urea concentrations in the native baseline region (0 - 1.8 M urea). The pD (uncorrected meter reading) was chosen to mimic the pH employed in a host of previous thermodynamic and kinetic folding experiments on α TS to define its folding free energy surface.¹⁶⁻¹⁹ Aggregation of the sample precluded the collection of data at and above 2 M urea. The requirement to prepare the sample and collect HSQC transients for ~30 minutes to obtain S/N sufficient for the precise calculation of the volume of the cross peaks, however, restricted the analysis to amide hydrogens whose protection factors over their solvent-exposed counterparts exceeded ~10⁵ under these conditions.²¹ This experimental limitation focused attention on those amide hydrogens that exchange through the I1 intermediate or higher energy partially-folded states.

The exchange properties of the amide hydrogens fall into four categories. The fast exchange category of amide hydrogens, corresponding to those found in side chains as well as in main chains located in loops and the ends of helices, are completely exchanged within the first ~30 min (Figure 2B). The second, intermediate exchange category of about 80 amide hydrogens,

largely contained in α -helices, then exchange over the course of the first few hours with observed rate constants on the order of 10^{-5} s^{-1} . The third, slow exchange category includes ~ 60 amide hydrogens, many of which are contained in β -strands (Figure 2C). This set of amides exchanges over a period of days to weeks and displays rate constants in the range from 10^{-6} to 10^{-8} s^{-1} . Finally, exchange over a period of months reveals an extremely stable core of 10 amide hydrogens (Figure 2D), the ultra-slow category, with estimated rate constants of the order of $\leq 10^{-9} \text{ s}^{-1}$. Unfortunately, the poor dispersion of cross peaks in a highly-helical 29 kDa protein at 600 MHz limited the number whose intensities could be measured reliably to monitor the protection against exchange. However, the measurable retardation of exchange for several dozen amide hydrogens, located in multiple elements of secondary structure distributed across the entire sequence of α TS, provided the probes required to assess the structures of high-energy partially-folded states.

The observed rates of exchange for 29 amide hydrogens at a series of urea concentrations between 0 and 1.8 M urea were obtained by fitting the fractional occupancy of hydrogen as a function of time to a single exponential function. Representative traces and fits at 0 M urea for representative amide hydrogens are shown in Figure 3A and 3B for two time regimes, hours and days, spanned by these data. The protection factors, i.e., the ratio of the observed exchange rate constant and the rate constant for a solvent-exposed amide hydrogen, $K_p = k_{\text{obs}}/k_{\text{ex}}$, range from 1.6×10^{-5} for the more rapidly exchanging amide hydrogens to 2.4×10^{-9} for several of the most strongly protected amide hydrogens in the absence of denaturant.

The interpretation of these data in terms of the thermodynamic or kinetic properties of the unfolding reactions enabling exchange depends on whether refolding from an exchange-competent state is faster or slower than exchange of the amide hydrogen with deuterium in the solvent-exposed state.^{22,23} If refolding is faster than exchange, the EX2 mechanism is operative and the observed rate constant is reduced by a factor that is proportional to the free energy difference between the protected and unprotected states. If refolding is slower than exchange, the EX1 mechanism is operative and the observed rate constant is reduced by a factor that is equal to the unfolding rate constant. Although HX under native-favoring conditions generally follows an EX2 mechanism,²⁴ this presumption was tested for α TS by repeating the exchange reaction at pD 8.2 and 1.2 M urea. The rate of exchange for all of the observed amide hydrogens was accelerated by a factor of about 2.5 compared to pD 7.8 and 1.2 M urea (data not shown), precisely the acceleration expected for the EX2 mechanism. Thus, the retardation in exchange for the slowly-exchanging amide hydrogens can be interpreted in terms of the stabilities of the corresponding partially-folded states relative to the native state of α TS, ΔG_{HX} .

A plot of the ΔG_{HX} values for these 29 amide hydrogens at a series of urea concentrations in the native baseline region is shown in Figure 4. The ΔG_{HX} values generally decrease monotonically with increasing urea concentration for all of the observed amide hydrogens up to 1.8 M urea. The retardation of HX for 12 amide hydrogens in $\alpha 1$, $\beta 2$ and $\beta 3$ is too great to allow reliable measurements of their ΔG_{HX} values below 0.8 M urea. Their ΔG^0_{HX} values in the absence of denaturant were estimated by linear extrapolation of the data at and above 0.8 M urea where reliable measurements can be made. Pace and colleagues²⁵ have previously shown that the free energy difference between two thermodynamic states is well-described by a linear dependence on urea concentration. Although local unfolding reactions could enhance exchange at lower denaturant concentration²⁴ for the amide hydrogens in this ultra-slow category, the distinct linear urea dependence above 0.8 M urea reveals the underlying large-scale unfolding reaction. Within the estimated errors, the decrease in ΔG_{HX} is linear with increasing urea concentration for all of the remaining amide hydrogens monitored.

The ΔG^0_{HX} values in the absence of denaturant range from 8.0 to 11.7 kcal mol⁻¹, and the m values, the urea-dependence of ΔG_{HX} , range from 0.6 to 2.0 kcal mol⁻¹ M⁻¹ (Table 1). Although the protection experienced by several amide hydrogens, D225 ($\alpha 7$), I240 ($\alpha 8'$), V121 ($\alpha 3$), G17 (turn before $\beta 1$), E42 ($\alpha 1$), was too weak to persist in even the lowest concentrations of urea, their protection factors in the absence of denaturant could be estimated from the retardation of exchange at 0 M urea (Figure 3A). The correlation of the estimates of ΔG^0_{HX} in the absence of denaturant with the position of the respective amide hydrogens in the β -barrel is shown in 2D format in Figure 5A and, including the α -helical amide hydrogens, in 3D format in Figure 5B.

As is evident in these complementary displays of the barrel architecture, the levels of protection against exchange are not distributed uniformly over all of the 8 β -strands and several of the α -helices. Several striking features emerge from the data displayed in Figures 5A and 5B and Table 1:

1. The most stable region is comprised of $\alpha 1$, $\beta 2$ and $\beta 3$, adjacent elements of secondary structure in the TIM barrel. Amide hydrogens between $\beta 2$ (L48 and L50) and $\beta 3$ (G98 and L100) and those protruding from $\beta 2$ (E49 and G51) and $\beta 3$ (I97, L99 and M101) and hydrogen-bonded to $\beta 1$ and $\beta 4$, respectively, in the native barrel are all exceedingly slow to exchange and have estimated ΔG^0_{HX} values between 10.5 and 11.7 kcal mol⁻¹ (Figure 4, violet symbols). The amide hydrogens in $\alpha 1$ are those associated with large nonpolar side chains, I37, L40 and I41, that define the hydrophobic face of the helix that docks on $\beta 1$ and $\beta 2$.
2. The next most stable region, $9.5 \leq \Delta G^0_{\text{HX}} \leq 10.5$ kcal mol⁻¹ (Figure 4, blue symbols), involves amide hydrogens at or near the N-termini of $\beta 1$ (A18 and F19), $\beta 4$ (S125) and $\beta 5$ (A149). Equivalent positions in $\beta 7$ and $\beta 8$ either do not have an amide hydrogen, P208 in $\beta 7$, or appear to exchange through a local unfolding reaction, G230 in $\beta 8$. The latter conclusion reflects the reduced dependence of ΔG_{HX} on the urea concentration²⁶ for G230 (see point 3 below).
3. The next level of protection, $8.5 \leq \Delta G^0_{\text{HX}} \leq 9.5$ kcal mol⁻¹ (Figure 4, green symbols), was offered by residues that are largely found in the central positions of these strands, e.g., V20 and L25 in $\beta 1$, L127 in $\beta 4$, I151 in $\beta 5$, and Y175 and L176 in $\beta 6$, or near the N-terminus of $\beta 8$, G230 and A231. The reduced values of m_{HX} for L25, G230 and A231 (Table 1) argue that their amide hydrogens exchange through local unfolding reactions.²⁶

Note that the amide hydrogen for G230 does not form a hydrogen bond with a carbonyl oxygen in $\beta 7$. Rather, the likely source of protection could be a buried water molecule that intervenes between $\beta 7$ and $\beta 8$. The Ramachandran angles, ϕ ψ for G230, +148, -171,¹⁴ are unique to glycine residues and enable a sharp turn in the chain that excludes bulk water from the amide linkage.

4. The next level of protection, $7.5 \leq \Delta G^0_{\text{HX}} \leq 8.5$ kcal mol⁻¹ (Figure 4, yellow symbols), was found for the only protected amide hydrogen in $\beta 7$, L209, another in $\beta 8$, I232, which is H-bonded to $\beta 7$, and M262 in $\alpha 8$.
5. The lowest measurable protection, $6.5 \leq \Delta G^0_{\text{HX}} \leq 8.0$ kcal mol⁻¹, was found for G17 in the turn before $\beta 1$, E42 in $\alpha 1$, V121 in $\alpha 3$, D225 in $\alpha 7$, and I240 in $\alpha 8'$.
6. The protection decreases ~monotonically from $\beta 2$ and $\beta 3$ towards both the N- and C-terminus of the sequence; amide hydrogens in $\beta 7$ and $\beta 8$ offer the least protection against exchange.
7. The protection is stronger at the N-terminus of $\beta 1$, $\beta 4$, $\beta 5$ and $\beta 8$ than at the C-terminus

8. The urea dependence of the ΔG°_{HX} values, i.e., the magnitude of m_{HX} , for the strongly-protected amide hydrogens increases linearly, within the estimated errors, with ΔG°_{HX} (Figure 6). This behavior implies that exchange occurs through a series of states of progressively increasing solvent-exposure of buried surface area and increasing free energy relative to the state that offers protection.²⁷
9. A subset of the most slowly exchanging amides in the most stable core region, $\alpha 1$, $\beta 2$ and $\beta 3$, display an unusually high protection at 1.6 M urea: I36 and L40 in $\alpha 1$, L48, in $\beta 2$ and I97, L99, L100 and M101 in $\beta 3$. By contrast, A47, E49 and G51 in $\beta 2$ and G98 in $\beta 3$ and other residues elsewhere in the protein display protection that varies linearly with urea concentration over the entire range examined. This observation rules out a systematic error in the urea concentration, the pH or the temperature for the peculiar behavior of the amide hydrogens associated with the most stable core. It is noteworthy that these large aliphatic nonpolar side chains are clustered together in the native structure. Perhaps a localized conformational change is triggered at this moderate urea concentration. The results of previous studies of the optical properties of αTS in the native baseline region suggest subtle conformational changes between 1.5 and 2.0 M urea (O. Bilsel, unpublished results).

The complexities of the protection patterns and the associated free energies of the species offering protection against exchange in αTS provide detailed insights into the structures of partially-folded states in equilibrium with the native conformation in the absence of denaturant. The relationship of these partially-folded states to folding intermediates detected by classical denaturation methods is considered in the Discussion section.

1H - ^{15}N HSQC analysis of ^{15}N -labeled Ile, Leu, Phe or Val αTS

These insights into the secondary structure of partially-folded states in αTS were complemented with information on the nonpolar side chains stabilizing I2, a highly-denatured intermediate known to be devoid of secondary structure.¹⁷ This goal was achieved by uniformly labeling αTS with either ^{15}N -leucine, ^{15}N -isoleucine, ^{15}N -phenylalanine or ^{15}N -valine in the appropriate auxotrophic strain of the *E. coli* host. The 1H - ^{15}N HSQC spectra of αTS containing selectively labeled ^{15}N -leucine and isoleucine residues were collected under conditions favoring the I2 and U states at 5 M and 8 M urea, respectively, and are shown in Figure 7. Although all 27 leucine and 19 isoleucine cross peaks are evident at 8 M urea where αTS is fully-denatured (Figures 7A and 7C, respectively), 4 leucine and 3 isoleucine cross peaks are absent in the 5 M urea spectra (Figures 7B and 7D, respectively). Comparable spectra of the ^{15}N -labeled Phe or Val αTS show the complete complement of cross peaks at both 5 and 8 M urea (data not shown). The failure to detect the entire set of leucine and isoleucine residues at 5 M urea implies that a subset of these large nonpolar side chains is dynamically-broadened by participation in the hydrophobic cluster known to stabilize I2.^{16,19}

To obtain further insights into the side chains that participate in this cluster, a double-labeling NMR experiment was performed on αTS . A previous 1D 1H NMR analysis had demonstrated that a number of side chains in the sequence corresponding to $\beta 3$ participate in the nonrandom structure defining the I2 species.¹⁹ The possibility that L99 in $\beta 3$ is one of the dynamically-broadened HSQC cross peaks at 5 M urea was tested by obtaining its assignment in the I2 species by double-labeling the unique G98-L99 peptide bond with ^{15}N -Leu and $^{13}C=O$ -Gly, respectively, using the auxotrophic strain DL39/DE3. The 2D HNC O NMR experiment, which provides the correlation between the backbone atoms 1H and ^{15}N of residue, 'i' and the $^{13}C=O$ of its preceding residue, revealed two cross peaks in 8 M urea (Figure 8). The more intense cross peak was assigned to the G98 $^{13}C=O$ - L99 NH, and the less intense peak to the overlap of two Ser-Leu peptide bonds, S6-L7 and S33-L34. Glycine and serine reside in the same biosynthetic pathway, providing an opportunity for the partial incorporation of the ^{13}C -label in

glycine into the carbonyl of serine.²⁸ The L99 assignment can be mapped onto the HSQC spectrum of α TS in 8 M urea, where L99 is seen to be one of the peaks that is dynamically broadened at 5 M urea (Figure 7B).

The correspondence between one of the elements of secondary structure that is most resistant to HX, β 3 (including L99), and one component of the nonrandom structure defining I2, L99, is striking and provides a crucial insight into the order of events in the equilibrium folding reaction of α TS (see Discussion).

DISCUSSION

Structural implications for folding intermediates in α TS

The structures of rare partially-folded states of α TS in equilibrium with the native state at low denaturant concentrations may, in principle, be unrelated to folding intermediates populated under mildly or strongly denaturing conditions. However, it is commonly accepted that the properties of discrete thermodynamic states in the absence of urea, can be obtained by simple extrapolation from unfolding reactions induced by molar concentrations of denaturant.²⁵ Thus, it is reasonable to presume that the native-state HX data and the ^{15}N -Leu and ^{15}N -Ile HSQC data can be interpreted in terms of the structures of folding intermediates of α TS predicted from a comprehensive set of thermodynamic and kinetic folding experiments (Figure 1).^{17, 18} The validity of this presumption is supported by the good agreement between the range of observed protection factors for the partially-folded states, spanning $\Delta G^{\circ}_{\text{HX}}$ values from 6.5 kcal mol⁻¹ to 11.7 kcal mol⁻¹ (Table 1), and the range previously determined for the I1, I_{BP} and I2 states, 6.2 to 11.2 kcal mol⁻¹ (Figure 1).

For the amide hydrogens that display fast exchange behavior (amides that exchange within the first ~30 minutes and that have estimated rate constants of $>10^{-5} \text{ s}^{-1}$), HX is expected to reflect primarily the properties of the native state. This supposition is based upon the fact that the native state is dominant under these conditions and the fact that the half-life of the native state under these conditions is ~40 minutes.¹⁸ Side chain amide hydrogens and main chain amide hydrogens in loops and some turns are observed to be fully exchanged in this time frame (Figure 2B). By contrast, amide hydrogens in α -helices and β -strands are largely protected in the N state. The inability to directly monitor the rates of exchange and protection factors for the rapidly-exchanging amide hydrogens also precludes the possibility of observing local unfolding reactions that would correspond to a set of microstates in the native manifold of conformers.

For the intermediate exchange category (exchange rate constants between 10^{-5} s^{-1} and 10^{-6} s^{-1}), additional amide hydrogens in loops and most of those in helices exchange with solvent; the β -strands remain protected. The estimated $\Delta G^{\circ}_{\text{HX}}$ values are ~7 kcal mol⁻¹. For example, amide hydrogens for E42 (α 1), V121 (α 3), D225 (α 7) and I240 (α 8') (Table 1 and Figure 3A) have $\Delta G^{\circ}_{\text{HX}}$ values between 6.5 and 7.5 kcal mol⁻¹. This range lies between the estimated free energies of the I1 and I_{BP} states (Figure 1), suggesting that HX occurs through microstates within the I1 manifold. The structural model implied for I1 has surface α -helices and turns preferentially frayed or unfolded while the β -strand network is intact. This interpretation for the I1 intermediate is consistent with the preferential loss of helical CD signal previously reported for the I1 intermediate at 3 M urea.¹⁷

17 well-resolved amide hydrogens belong to the slow exchange category (rate constants between 10^{-6} s^{-1} and 10^{-8} s^{-1}) and exchange with $\Delta G^{\circ}_{\text{HX}}$ values between 7.9 and 9.9 kcal mol⁻¹ (Table 1 and Figure 4). These amide hydrogens, which are located in α 0, a loop before β 1, β 1, α 1, β 4, β 5, β 6, β 7, β 8 and α 8, must be exchanging through microstates in the I1 or I_{BP} manifolds. If HX occurred solely through the I2 state, common $\Delta G^{\circ}_{\text{HX}}$ values of 11.2 kcal

mol^{-1} would have been observed. The observation that the β -strands begin to exchange precisely at the same free energy as that previously determined for the I_{BP} state (Figure 9)¹⁸ strongly suggests that HX is occurring through an ensemble of microstates in I_{BP} . The putative I_{BP} thermodynamic state is comprised of an ensemble of weakly-coupled microstates whose most stable core is comprised of side chains from $\alpha 1$, $\beta 2$ and $\beta 3$. A similar linear relationship between ΔG°_{HX} and m_{HX} for a partially-folded state in RNase H, an $\alpha + \beta$ motif, was also interpreted to represent a continuum of microstates.²⁷

The pattern of protection in both $\beta 2$ and $\beta 3$, strings of 5 consecutive amide hydrogens, is difficult to understand in terms of the β -barrel architecture because inter-strand amide hydrogens in $\beta 1$ and $\beta 4$ are not as strongly protected. Although one could argue for solvent exclusion mechanisms,^{22,29,30} another intriguing possibility is that both segments initially form α -helices.²² A cluster of three helices, two of which represent non-native structures for $\beta 2$ and $\beta 3$, might provide a platform for the subsequent recruitment of weakly-coupled $\beta\alpha$ -hairpins to form the ensemble of microstates defining I_{BP} . The proposal for an ensemble of weakly-coupled states for I_{BP} is also consistent with the reduced m -value but similar ellipticity compared to the I_1 state. Non-native α helices observed in the I_{BP} species of β -lactoglobulin^{31,32} and in the molten globule state of canine milk lysozyme³³ show that transient helical structure can precede β strand formation.

Another striking observation in the native-state exchange data is the trend of enhanced protection at the N-terminus of several β strands compared to the C-terminus (Figure 5A). The asymmetric behavior signifies a prominent role for the N-termini of the β strands in stabilizing the nascent β -barrel. This observation is consistent with a similar conclusion derived from a protein engineering analysis of the phosphoribosylanthranilate isomerase TIM barrel³⁴ and with the speculation that the loops at C-termini of strands must be sufficiently flexible to enable catalytic function.³⁵ Intriguingly, the C-termini of the helices in αTS are connected to the N-termini of the subsequent strands by short loops/turns that contain an average of 4 residues (range: 3 to 5). These tight loops in general also contain residues that show a propensity to favor turn conformations, such as Gly, Asn, Asp and Ala. In contrast, the loops linking the C-termini of the strands to the subsequent helices contain an average of 8 residues (range: 5 to 10) and are responsible for substrate binding. The decreased loop entropy penalty and the increased turn propensity of the $\alpha\beta$ loops compared to their $\beta\alpha$ counterparts are both likely to contribute to the enhanced protection observed for the N-termini of the β -strands.

The 13 most strongly protected amide hydrogens, located in $\alpha 1$ (I37 and L40), $\beta 2$ (A47, L48, E49, L50 and G51), $\beta 3$ (I97, G98, L99, L100 and M101) and $\beta 4$ (S125), exchange with rate constants $\leq 10^{-8} \text{ s}^{-1}$ in the absence of denaturant. In some cases, their ΔG°_{HX} values in buffer were estimated by extrapolation of the values at and above 1.2 M urea (Figure 4) and found to range from 10.2 to 11.5 kcal mol^{-1} , with an error of $\sim 0.5 \text{ kcal mol}^{-1}$. These amide hydrogens exchange either through microstates in the upper reaches of the I_1 and I_{BP} manifolds or through the I_2 intermediate. The absence of protection with ΔG°_{HX} values greater than $11.7 \pm 0.4 \text{ kcal mol}^{-1}$, within error of the estimated relative stability of the I_2 state vs. the N state, $11.2 \pm 0.2 \text{ kcal mol}^{-1}$, is consistent with an I_2 state not possessing a stable hydrogen bond network that can offer protection against exchange in the absence of denaturant. This conclusion, however, does not preclude the possibility of weakly-folded helices that could contribute to the far-UV CD signal but not to protection at low urea concentrations.

As shown in Figure 7, a subset of 4 leucine and 3 isoleucine side chains, including L99, is responsible for stabilizing the I_2 state at 5 M urea. The participation of L99 in the exchange-resistant core in the absence of denaturant and in the non-random structure for I_2 observed at 5 M urea cannot be coincidental. The observation implies that most of the leucine and isoleucine side chains strongly protecting amide hydrogens in $\alpha 1$ $\beta 2$ and $\beta 3$, a total of 5 leucines and 2

isoleucines, might serve as the core of the hydrophobic cluster thought to define I2.^{16,19} It has previously been suggested that branched aliphatic side chains amino acids may play a crucial role in defining the hydrophobic clusters responsible for stabilizing the I_{BP} species in α TS.³⁶⁻³⁸ Consistent with the **B**ranched **A**liphatic **S**ide **C**hain (BASiC) hypothesis is the selective elimination of I_{BP} by alanine replacements for a set of leucine and isoleucine residues that reside in the most stable core identified by the native state HX experiments, α 1, β 2 and β 3.^{36,37}

Resolution of alternative views of the structure of the I1 intermediate

The conclusion that the I1 intermediate has an intact β -barrel with frayed helices contrasts with previous views of its structure. Fragmentation analysis has shown that the urea-induced CD-detected unfolding transition for the 1-188 fragment, $\alpha_0(\beta/\alpha)_{1-6}$, and the I1 [a2opleftrt] I2/U transition for the full-length α TS are nearly coincident.^{39,40} The similar ellipticity and stability properties of the fragment and the I1 intermediate were interpreted to mean that the structure of I1 is primarily dictated by the behavior of this large N-terminal segment of the protein. A HX-MS analysis of protection of main chain amide hydrogens against exchange with solvent by the I1 species showed that >50% protection against a 5 s pulse of ²H₂O at pH 7.8 and 25 °C in 3 M urea was only displayed by the $(\beta/\alpha)_{1-4}$ segment of α TS.⁴¹ These apparently disparate views of I1 can be reconciled with the HX-NMR results to provide a more accurate picture of its structure.

The fragmentation analysis depended upon the global measurement of the secondary structure via the far-UV ellipticity. The bias towards the stronger helix signal means that these measurements are inherently less sensitive to the signal from the β -barrel and, therefore, are not a reliable probe of its integrity. The fragmentation analysis is further complicated by the presence of two-slowly interconverting folded species in the absence of denaturant, a situation that makes the measurements of its stability and estimates of its structure ambiguous. The stronger protection against HX for $(\beta/\alpha)_{1-4}$ assigned to the I1 species in the HX-MS experiment is likely a reflection of the ensemble of microstates present at equilibrium under these conditions. The pulse intensity in the mass-labeling experiment was sufficiently strong to cause significant exchange for amide hydrogens in β 5- β 8, but not in β 1- β 4. Inspection of the HX-NMR data in Figure 5A shows that β 1- β 4 are distinct from β 5- β 8 in that the former group has amide hydrogens with $\Delta G^\circ_{\text{HX}} > 9.5 \text{ kcal mol}^{-1}$ and the latter group does not. Thus, the discrepancies in these various estimates of the structure of I1 reflect the ambiguities of fragmentation analysis and of pulse-labeling under moderate urea concentrations when multiple partially-folded forms are present. The native-state HX experiment on the full-length protein, under strongly folding conditions that enable the discrimination of the structures of I1 and I_{BP}, states provides evidence that I1 has an intact β -barrel under conditions favoring the native conformation.

Development of structure in the equilibrium folding mechanism of α TS

The close correspondence of the partially-folded states detected by the native state HX-NMR experiment and the intermediates detected by classical denaturation experiments enables a description of the development of structure during the equilibrium folding reaction for α TS. At high denaturant concentrations, evidence available from small angle x-ray scattering experiments⁴² (Y. Wu, unpublished results) supports a random coil conformation for the U state. As the urea concentration is decreased, a hydrophobic cluster forms around a nucleus of branched aliphatic side chains from segments corresponding to α 1, β 2 and β 3. It has previously been shown that other nonpolar segments, distal in sequence, also participate in this hydrophobic cluster.^{36,37} The lack of HX protection in the absence of denaturant for I2 demonstrate that it has minimal, if any stable secondary structure. Thus, I2, is stabilized by the hydrophobic effect and not by hydrogen bonding.

As the urea concentration is further decreased, the I1 and I_{BP} states become populated. Although the far-UV CD signals of these two intermediates are comparable, their protection patterns against exchange are quite different. The I1 species has an intact β -barrel with frayed and/or unfolded helices, which transiently associate with the nonpolar exterior of the β -barrel. By contrast, the I_{BP} species initially forms around what may be a trio of helices, two of which (those which ultimately form β 2 and β 3 in the native TIM barrel) would correspond to non-native structures stabilized by nonpolar interactions between branched aliphatic side chains. The linear and direct relationship between the $\Delta G^{\circ}_{\text{HX}}$ and m-values for the amide hydrogens in the β -barrel for α TS (Figure 6) is consistent with the progressive development of structure and stability towards both termini in I_{BP}. The asymmetric protection pattern in several of the β -strands, however, is not consistent with the formation of individual and complete elements of secondary structure. By contrast, the exchange properties of the all-helical cytochrome c have been interpreted in terms of the unfolding of discrete elements of secondary structure.²⁶ Further HX-NMR experiments on the properties of folding intermediates for other motifs are required to test the generality of either of these observations.

As the urea concentration is further reduced, the native state dominates the population of α TS. The N state is distinguished from both the I1 and I_{BP} states by the establishment of a stable helical shell around the β -barrel framework. Complementary near-UV CD, absorbance and fluorescence data,¹⁷ show that this transition is also accompanied by the formation of chiral, solvent-excluded environments found for the aromatic side chains in the native tertiary structure.

Perspective

The native-state HX-NMR experiment has provided detailed insights into the secondary structures of the partially-folded states that exist in equilibrium with the native state of α TS at low concentrations of urea. Complementary HSQC experiments on a highly-denatured state revealed a significant role for isoleucine and leucine side chains in a hydrophobic cluster that lacks stable secondary structure, even in the absence of denaturant. A previous comprehensive analysis of the folding mechanism of α TS has shown that folding intermediates with very similar free energies dominate the population of protein during the urea-induced equilibrium folding reaction.¹⁸ Thus, the native-state HX experiment not only provided information on the structures and energetics of partially-folded states in the absence of denaturant but also on the development of structure and stability during the equilibrium folding reaction. Closely-related equilibrium folding mechanisms for two other TIM barrel proteins of low sequence identity^{38,43,44} suggest that the multi-step development of structure and stability observed for α TS may be a common feature of the TIM barrel motif.

Materials and Methods

Protein Preparation and Purification

Uniformly ¹⁵N-labeled α TS was obtained by growing *E. coli* strain BL21(DE3) harboring the plasmid pT7.WS1 in M9 medium containing 6 g/l of Na₂HPO₄, 3 g/l of KH₂PO₄, 0.5 g/l of NaCl and 2 g/l glucose. The protein was expressed and purified according to the published procedure.²⁰ Auxotrophic strain DL39avtA/DE3 was used to obtain α TS selectively ¹⁵N-labeled at Leu and Ile residues, and the same strain was also used to obtain the ¹⁵N-Leu / ¹³C=O-Gly labeled α TS. For selective labeling, the cells were grown in M9 medium supplemented with nucleic acid bases and all unlabeled amino acids except for ¹⁵N-Leu and ¹³C=O-Gly. The ¹⁵N and ¹³C enriched amino acids were obtained from Cambridge Isotopes and Deuterium oxide (99.9%) from Sigma Aldrich.

Hydrogen Deuterium (H/D) Exchange

For the H/D exchange experiments, purified α TS was dialyzed in aqueous buffer containing 50 mM potassium phosphate, pH 7.8, 0.2 mM EDTA and 2 mM BME and then lyophilized. Exchange was initiated by dissolving the lyophilized protein in deuterium oxide buffer containing concentrations of d_4 -urea ranging from 0 to 1.8 M. Upon addition of deuterium oxide the samples were immediately transferred to NMR tubes and placed in the spectrometer. Time between initiation of exchange, transferring to the NMR tube, placing in the spectrometer, tuning and shimming, filtering, and the beginning of data collection averaged 10 min. The protein concentration was ~ 0.4 mM and the H/D exchange was monitored at 25 °C, and when outside of the spectrometer the samples were kept in a temperature-controlled water bath at 25 °C. The spectra were staggered in time over periods of hours to days. The hydrogen-deuterium exchange was monitored by acquiring successive ^1H - ^{15}N HSQC spectra with 4-64 transients per increment and typically 128 increments were acquired. The spectral widths were 10000 Hz in the ^1H dimension and 2200 in the ^{15}N dimension. All NMR experiments were recorded on a Varian 600 MHz spectrometer.

Data Analysis

Hydrogen exchange rates were obtained from the decay of ^1H - ^{15}N HSQC cross peak intensities as a function of exchange time, defined as the period of time from the dissolution of lyophilized α TS in $^2\text{H}_2\text{O}$ buffer, to the end of the each ^1H - ^{15}N HSQC experiment. Spectral processing and cross peak intensities were obtained using FELIX 97.0 (Accelrys). Peak intensity as a function of time was fit to a single exponential decay, $I = I_0 \exp(-k_{\text{obs}} * t)$, with the initial intensity I_0 and the observed exchange rate k_{obs} as free variables in the fit. The uncertainties for the k_{obs} values were taken as standard errors of the fits. The $\Delta G^{\circ}_{\text{HX}}$ was obtained from the equation: $\Delta G^{\circ}_{\text{HX}} = -RT \ln(k_{\text{obs}}/k_{\text{int}})$, where k_{int} is the intrinsic exchange rate calculated for amide protons in unstructured peptides. The k_{int} values were obtained using the program SPHERE (<http://www.fccc.edu/research/labs/roder/sphere/>).

Acknowledgements

The authors wish to thank Drs. Jill Ann Zitzewitz, Osman Bilse and Zhenyu Gu for critical reviews of the manuscript and many helpful comments. Initial help by Dr. Piotr Dobrowolski with the Varian NMR instrument is gratefully acknowledged. This work was supported by the National Institutes of Health through grant GM23303 to CRM.

Abbreviations

α TS, alpha subunit of tryptophan synthase
 CD, circular dichroism
 $\Delta G^{\circ}_{\text{HX}}$, free energy of protection against hydrogen exchange in the absence of denaturant
 HSQC, Heteronuclear Single Quantum Coherence
 HX, hydrogen-to-deuterium exchange of amide hydrogens
 I1, equilibrium unfolding intermediate of α TS highly populated at 3 M urea
 I2, equilibrium unfolding intermediate of α TS highly populated at 5 M urea
 I_{BP}, kinetic intermediate of α TS populated within the stopped-flow burst phase (<5 ms)
 N, native state
 NMR, Nuclear Magnetic Resonance
 TIM, triosephosphate isomerase
 U, unfolded state

References

1. Tanford C. Protein denaturation. Adv. Protein Chem 1968;23:121–282. [PubMed: 4882248]
2. Jackson SE. How do small single-domain proteins fold? Fold. Des 1998;3:R81–R91. [PubMed: 9710577]

3. Bilsel O, Matthews CR. Barriers in protein folding reactions. *Adv. Protein Chem* 2000;53:153–207. [PubMed: 10751945]
4. Kamagata K, Arai M, Kuwajima K. Unification of the folding mechanisms of non-two-state and two-state proteins. *J. Mol. Biol* 2004;339:951–965. [PubMed: 15165862]
5. Cecconi C, Shank EA, Bustamante C, Marqusee S. Direct observation of the three-state folding of a single protein molecule. *Science* 2005;309:2057–2060. [PubMed: 16179479]
6. Raschke TM, Marqusee S. The kinetic folding intermediate of ribonuclease H resembles the molten globule and partially unfolded molecules detected under native conditions. *Nat. Struct. Biol* 1997;4:298–304. [PubMed: 9095198]
7. Feng H, Zhou Z, Bai Y. A protein folding pathway with multiple folding intermediates at atomic resolution. *Proc. Natl. Acad. Sci. U. S. A* 2005;102:5026–5031. [PubMed: 15793003]
8. Chamberlain AK, Handel TM, Marqusee S. Detection of rare partially folded molecules in equilibrium with the native conformation of RNaseH. *Nat. Struct. Biol* 1996;3:782–787. [PubMed: 8784352]
9. Chamberlain AK, Marqusee S. Touring the landscapes: partially folded proteins examined by hydrogen exchange. *Structure* 1997;5:859–863. [PubMed: 9261079]
10. Bai Y. Energy barriers, cooperativity, and hidden intermediates in the folding of small proteins. *Biochem. Biophys. Res. Commun* 2006;340:976–983. [PubMed: 16405866]
11. Rumbley J, Hoang L, Mayne L, Englander SW. An amino acid code for protein folding. *Proc. Natl. Acad. Sci. U. S. A* 2001;98:105–112. [PubMed: 11136249]
12. Orengo CA, Michie AD, Jones S, Jones DT, Swindells MB, Thornton JM. CATH--a hierarchic classification of protein domain structures. *Structure* 1997;5:1093–1108. [PubMed: 9309224]
13. Nagano N, Orengo CA, Thornton JM. One fold with many functions: the evolutionary relationships between TIM barrel families based on their sequences, structures and functions. *J. Mol. Biol* 2002;321:741–765. [PubMed: 12206759]
14. Hyde CC, Ahmed SA, Padlan EA, Miles EW, Davies DR. Three-dimensional structure of the tryptophan synthase alpha 2 beta 2 multienzyme complex from *Salmonella typhimurium*. *J. Biol. Chem* 1988;263:17857–17871. [PubMed: 3053720]
15. Matthews CR, Crisanti MM, Manz JT, Gepner GL. Effect of a single amino acid substitution on the folding of the alpha subunit of tryptophan synthase. *Biochemistry* 1983;22:1445–1452. [PubMed: 6132619]
16. Saab-Rincon G, Froebe CL, Matthews CR. Urea-induced unfolding of the alpha subunit of tryptophan synthase: one-dimensional proton NMR evidence for residual structure near histidine-92 at high denaturant concentration. *Biochemistry* 1993;32:13981–13990. [PubMed: 8268176]
17. Gualfetti PJ, Bilsel O, Matthews CR. The progressive development of structure and stability during the equilibrium folding of the alpha subunit of tryptophan synthase from *Escherichia coli*. *Protein Sci* 1999;8:1623–1635. [PubMed: 10452606]
18. Bilsel O, Zitzewitz JA, Bowers KE, Matthews CR. Folding mechanism of the alpha-subunit of tryptophan synthase, an alpha/beta barrel protein: global analysis highlights the interconversion of multiple native, intermediate, and unfolded forms through parallel channels. *Biochemistry* 1999;38:1018–1029. [PubMed: 9893998]
19. Saab-Rincon G, Gualfetti PJ, Matthews CR. Mutagenic and thermodynamic analyses of residual structure in the alpha subunit of tryptophan synthase. *Biochemistry* 1996;35:1988–1994. [PubMed: 8639683]
20. Vadrevu R, Falzone CJ, Matthews CR. Partial NMR assignments and secondary structure mapping of the isolated alpha subunit of *Escherichia coli* tryptophan synthase, a 29-kD TIM barrel protein. *Protein Sci* 2003;12:185–191. [PubMed: 12493842]
21. Englander SW, Poulsen FM. Hydrogen-tritium exchange of the random chain polypeptide. *Biopolymers* 1969;7:379–393.
22. Englander SW, Kallenbach NR. Hydrogen exchange and structural dynamics of proteins and nucleic acids. *Q. Rev. Biophys* 1983;16:521–655. [PubMed: 6204354]
23. Bai Y. Kinetic evidence for an on-pathway intermediate in the folding of cytochrome c. *Proc. Natl. Acad. Sci. U S A* 1999;96:477–480. [PubMed: 9892658]
24. Englander SW, Sosnick TR, Englander JJ, Mayne L. Mechanisms and uses of hydrogen exchange. *Curr. Opin. Struct. Biol* 1996;6:18–23. [PubMed: 8696968]

25. Myers JK, Pace CN, Scholtz JM. Denaturant m values and heat capacity changes: Relation to changes in accessible surface areas of protein unfolding. *Protein Sci* 1995;4:2138–2148. [PubMed: 8535251]
26. Bai Y, Sosnick TR, Mayne L, Englander SW. Protein folding intermediates: native-state hydrogen exchange. *Science* 1995;269:192–197. [PubMed: 7618079]
27. Parker MJ, Marqusee S. A statistical appraisal of native state hydrogen exchange data: evidence for a burst phase continuum? *J. Mol. Biol* 2000;300:1361–1375. [PubMed: 10903874]
28. Waugh DS. Genetic tools for selective labeling of proteins with alpha-15N-amino acids. *J. Biomol. NMR* 1996;8:184–192. [PubMed: 8914274]
29. Alexandrescu AT, Jaravine VA, Dames SA, Lamour FP. NMR hydrogen exchange of the OB-fold protein LysN as a function of denaturant: the most conserved elements of structure are the most stable to unfolding. *J Mol Biol* 1999;289:1041–1054. [PubMed: 10369781]
30. Jaravine VA, Rathgeb-Szabo K, Alexandrescu AT. Microscopic stability of cold shock protein A examined by NMR native state hydrogen exchange as a function of urea and trimethylamine N-oxide. *Protein Sci* 2000;9:290–301. [PubMed: 10716181]
31. Hamada D, Segawa S, Goto Y. Non-native alpha-helical intermediate in the refolding of beta-lactoglobulin, a predominantly beta-sheet protein. *Nat. Struct. Biol* 1996;3:868–873. [PubMed: 8836104]
32. Forge V, Hoshino M, Kuwata K, Arai M, Kuwajima K, Batt CA, Goto Y. Is folding of beta-lactoglobulin non-hierarchic? Intermediate with native-like beta-sheet and non-native alpha-helix. *J. Mol. Biol* 2000;296:1039–1051. [PubMed: 10686102]
33. Watanabe M, Kobashigawa Y, Aizawa T, Demura M, Nitta K. A non-native alpha-helix is formed in the beta-sheet region of the molten globule state of canine milk lysozyme. *Protein J* 2004;23:335–342. [PubMed: 15328889]
34. Urfer R, Kirschner K. The importance of surface loops for stabilizing an eightfold beta alpha barrel protein. *Protein Sci* 1992;1:31–45. [PubMed: 1304881]
35. Branden, C.; Tooze, J. *Introduction to Protein Structure*. 2nd edit. Garland Science Publishing; 1999.
36. Wu Y, Vadrevu R, Yang X, Matthews CR. Specific structure appears at the N terminus in the sub-millisecond folding intermediate of the alpha subunit of tryptophan synthase, a TIM barrel protein. *J. Mol. Biol* 2005;351:445–452. [PubMed: 16023136]
37. Wu Y, Vadrevu R, Kathuria S, Yang X, Matthews CR. A tightly packed hydrophobic cluster directs the formation of an off-pathway sub-millisecond folding intermediate in the alpha subunit of tryptophan synthase, a TIM barrel protein. *J. Mol. Biol* 2007;366:1624–1638. [PubMed: 17222865]
38. Gu Z, Zitzewitz JA, Matthews CR. Mapping the structure of folding cores in TIM barrel proteins by hydrogen exchange mass spectrometry: the roles of motif and sequence for the indole-3-glycerol phosphate synthase from *Sulfolobus solfataricus*. *J. Mol. Biol* 2007;368:582–594. [PubMed: 17359995]
39. Higgins W, Fairwell T, Miles EW. An active proteolytic derivative of the alpha subunit of tryptophan synthase. Identification of the site of cleavage and characterization of the fragments. *Biochemistry* 1979;18:4827–4835. [PubMed: 389284]
40. Zitzewitz JA, Gualfetti PJ, Perkons IA, Wasta SA, Matthews CR. Identifying the structural boundaries of independent folding domains in the alpha subunit of tryptophan synthase, a beta/alpha barrel protein. *Protein Sci* 1999;8:1200–1209. [PubMed: 10386870]
41. Rojsajjakul T, Wintrode P, Vadrevu R, Robert Matthews C, Smith DL. Multi-state unfolding of the alpha subunit of tryptophan synthase, a TIM barrel protein: insights into the secondary structure of the stable equilibrium intermediates by hydrogen exchange mass spectrometry. *J. Mol. Biol* 2004;341:241–253. [PubMed: 15312776]
42. Gualfetti PJ, Iwakura M, Lee JC, Kihara H, Bilsel O, Zitzewitz JA, Matthews CR. Apparent radii of the native, stable intermediates and unfolded conformers of the alpha-subunit of tryptophan synthase from *E. coli*, a TIM barrel protein. *Biochemistry* 1999;38:13367–13378. [PubMed: 10529212]
43. Gu Z, Rao MK, Forsyth WR, Finke JM, Matthews CR. Structural analysis of kinetic folding intermediates for a TIM barrel protein, indole-3-glycerol phosphate synthase, by hydrogen exchange mass spectrometry and Go-model simulation. *J Mol. Biol.* 2007in press
44. Forsyth WR, Bilsel O, Gu ZY, Matthews CR. Topology and sequence in the folding of a TIM barrel protein: Global analysis highlights partitioning between transient off-pathway and stable on-pathway

- folding intermediates in the complex folding mechanism of a (beta/alpha)₈ barrel of unknown function from *B. subtilis*. *J. Mol. Biol.* 2007;372:236–253. [PubMed: 17619021]
45. DeLano, WL. The PyMOL molecular graphics system. DeLano Scientific; 2002.

Scheme 1.

Figure 1.

The reaction coordinate diagram of α TS displaying the relative free energies of the native state, N, the intermediate states, I1, I2, and I_{BP}, and the unfolded state, U in the absence of denaturant at pH 7.8 and 25 °C. The I_{BP} species is drawn to the right of the U species to indicate that I_{BP} is a misfolded species that must unfold and back-track through the I2 or U states to reach the productive folding pathway.

Figure 2.

Hydrogen to deuterium exchange in αTS . (A). The ^1H - ^{15}N HSQC spectrum of αTS in H_2O . ^1H - ^{15}N HSQC spectra of αTS in $^2\text{H}_2\text{O}$ after (B) 30 min, (C) 24 hours and (D) 200 days of exchange in the absence of denaturant at pH 7.8, 25 °C. Representative residues that exchange slowly and could be monitored in limited amounts of urea are labeled in (C). The residues that are still observed after 200 days of exchange are labeled in (D). The buffer contained 58 mM KPi, 0.2 mM EDTA and 1 mM DTE.

Figure 3. Time dependent decrease in the ^1H - ^{15}N cross peak intensity of the backbone amide hydrogens for representative set of residues that exchange in the time frame between (A) hours ($\blacktriangledown$ E42, $\blacksquare$ V121, $\circ$ D225, $\circ$ I240, and (B) days and weeks ($\blacktriangle$ F19, $\circ$ V20, $*$ L25, $\blacksquare$ A149, $\blacktriangledown$ G230) in 0 M urea. The peak intensities were normalized with respect to the intensity measured in the first spectrum. The continuous lines are the fits to a single exponential decay function.

Figure 4.

Profiles of ΔG_{HX} as a function of urea concentration at pH 7.8, 25 °C for the slowly-exchanging amide hydrogens located in α -helices and β -strands of α TS. The lines represent the linear fits required to obtain the m -value and ΔG_{HX}^0 in the absence of denaturant. The color scheme denotes the range of ΔG_{HX}^0 values: violet, 10.5 - 11.7 kcal mol⁻¹; blue, 9.5 - 10.5 kcal mol⁻¹; green, 8.5 - 9.5 kcal mol⁻¹; yellow, 7.5 - 8.5 kcal mol⁻¹. Accurate values of ΔG_{HX} for the most slowly exchanging amides (violet) could not be obtained at 0 and 0.4 M urea; dashed lines denote the extrapolated values for ΔG_{HX}^0 at 0 M urea.

Figure 5.

(A) 2D representation of the hydrogen bonding pattern of the β sheet in α TS. The arrows ($\rightarrow$) depict the hydrogen bonds, pointing from the backbone amide (NH) to the backbone carbonyl oxygen (C=O), with the exception of F19 in $\beta 1$ and I97 in $\beta 3$ where their respective backbone amides are hydrogen bonded to the side chain carboxyl of D46 and D124, respectively. The circles with a cross indicate proline residues, and filled circles indicate that the amide hydrogen is not involved in hydrogen bonding in the crystal structure. The empty circles indicate either the absence of the NMR assignment or the inability to obtain accurate fits due to spectral overlap. The color scheme is indicative of the decreasing ΔG^0_{HX} (from violet to yellow) in the absence of urea. (B) The measured hydrogen exchange free energies, 6.5 to 11.7 kcal mol⁻¹, in 0 M urea, mapped on the crystal structure of α TS. The figure is generated using Pymol⁴⁵ from the protein data bank file 1bks for α TS.¹⁴ The coordinates for α TS were extracted from the tetrameric tryptophan synthase, $\alpha_2\beta_2$, to generate the figure.

Figure 6.

The calculated free energy of exchange in the absence of urea, ΔG°_{HX} , plotted against m_{HX} , the denaturant dependence of the free energy of exchange for the amide hydrogens of α TS. The data were taken from Table 1.

Figure 7.

^1H - ^{15}N HSQC spectra of αTS selectively labeled at Leu and Ile residues. ^{15}N -Leu labeled αTS in (A) 8 M urea and in (B) 5 M urea. The cross peaks corresponding to all of the 27 leucines in αTS are observed in 8 M urea. The 4 missing cross peaks in 5 M urea are indicated by 'X'. ^{15}N -Ile labeled αTS in (C) 8 M and in (D) 5 M urea. All 19 isoleucine cross peaks are observed at 8 M urea, and the 3 missing cross peaks in 5 M urea are indicated by 'X'.

Figure 8.

The backbone $^{13}\text{C=O}$ resonance assignment of Gly98 of αTS . 2D HNCO spectrum of αTS selectively ^{15}N -labeled at Leu and $^{13}\text{C=O}$ -labeled at Gly residues in 8 M urea. The unique ^{98}Gly - ^{99}Leu pair in the sequence is identified from the observation of a single strong cross peak at the ^1H chemical shift at ~ 8 ppm and the $^{13}\text{C=O}$ chemical shift at 169 ppm. The other minor peak reflects the contribution from the two Ser-Leu pairs in the sequence.

Figure 9.

Relative stabilities of the native state, N, the partially folded states, I1, I_{BP}, I2 and the unfolded state, U, of α TS under native conditions. The dashed lines are a representation of the ensemble of microstates within the N, I1 and I_{BP} intermediates. The microstates for the N state could not be directly observed (experimental limitations), but are presumed to exist because a significant fraction of the amide hydrogens exchange out of the native conformation (see text).

Table 1Free energy for protection against exchange for amide hydrogens in α TS.^a

Residue	Structure ^b	$\Delta G^{\circ}_{\text{HX}}$ ^c (kcal mol ⁻¹)	m_{HX} ^c (kcal mol ⁻¹ M ⁻¹)
F8	$\alpha 0$	8.6(± 0.1)	0.84(± 0.1)
G17 ^d	loop	8.2	ND
A18	$\beta 1$	9.5(± 0.1)	1.0(± 0.1)
F19	$\beta 1$	9.6(± 0.2)	1.2(± 0.1)
V20	$\beta 1$	8.8(± 0.1)	1.1(± 0.1)
L25	$\beta 1$	8.7(± 0.1)	0.6(± 0.1)
I36	$\alpha 1$	9.9(± 0.2)	0.7(± 0.2)
I37	$\alpha 1$	11.7(± 0.6)	2.2(± 0.5)
L40	$\alpha 1$	10.6(± 0.4)	1.2(± 0.3)
E42 ^d	$\alpha 1$	7.2	ND
A47	$\beta 2$	11.7(± 0.4)	2.0(± 0.3)
L48	$\beta 2$	11.0(± 0.4)	2.0(± 0.3)
E49	$\beta 2$	10.6(± 0.3)	1.8(± 0.3)
L50	$\beta 2$	10.6(± 0.4)	1.9(± 0.3)
G51	$\beta 2$	11.3(± 0.2)	1.6(± 0.2)
I97	$\beta 3$	10.3(± 0.8)	1.7(± 0.5)
G98	$\beta 3$	11.7(± 0.4)	1.8(± 0.3)
L99	$\beta 3$	11.0(± 0.5)	1.7(± 0.3)
L100 ^e	$\beta 3$	≥ 10.3	ND
M101	$\beta 3$	10.4(± 0.3)	1.3(± 0.2)
V121 ^d	$\alpha 3$	6.9	ND
S125	$\beta 4$	10.2(± 0.2)	1.3(± 0.1)
L127	$\beta 4$	9.2(± 0.2)	1.3(± 0.2)
A149	$\beta 5$	9.7(± 0.1)	1.3(± 0.1)
I151	$\beta 5$	9.3(± 0.2)	1.6(± 0.2)
Y175	$\beta 6$	9.1(± 0.1)	0.9(± 0.1)
L176	$\beta 6$	8.7(± 0.1)	1.0(± 0.1)
L209	$\beta 7$	8.0(± 0.1)	0.8(± 0.1)
D225 ^d	$\alpha 7$	6.8	ND
G230	$\beta 8$	9.1(± 0.1)	0.7(± 0.1)
A231	$\beta 8$	9.1(± 0.1)	0.7(± 0.1)
I232	$\beta 8$	8.0(± 0.1)	0.7(± 0.1)
I240 ^d	$\alpha 8'$	6.8	ND
M262	$\alpha 8$	8.2(± 0.1)	0.6(± 0.1)

^a In the absence of urea at pH 7.8 and 25 °C; the buffer contained 50 mM potassium phosphate, 0.2 mM EDTA, 1 mM DTE.^b The location of the residues in the secondary structure.^c m_{HX} values and errors for $\Delta G^{\circ}_{\text{HX}}$ were obtained from fitting the protection data to $\Delta G_{\text{HX}} = \Delta G^{\circ}_{\text{HX}} - m_{\text{HX}}[\text{urea}]$.^d Estimates of ΔG_{HX} were based upon the measurement of exchange in the absence of urea; the m_{HX} values were not determined.^e $\Delta G^{\circ}_{\text{HX}}$ for L100 represents a minimum estimate. The scatter in the data prevented accurate determination.