Smith, D. D. S., & MacKenzie, R. E. (1983) *Can. J. Biochem. Cell Biol. 61*, 1166–1171.
Staben, C., & Rabinowitz, J. C. (1983) in *Chemistry and Biology of Pteridines* (Blair, J. A., Ed.) pp 381–385, Walter de Gruyter & Co., Berlin.
Tan, L. U. L., & MacKenzie, R. E. (1977) *Biochim. Biophys. Acta 485*, 52–59.
Tan, L. U. L., Drury, E. J., & MacKenzie, R. E. (1977) *J. Biol. Chem. 252*, 1117–1122.
Uyeda, K., & Rabinowitz, J. C. (1967) *J. Biol. Chem. 242*, 4378–4385.
Wasserman, G. F., Benkovic, P. A., Young, M., & Benkovic, S. J. (1983) *Biochemistry 22*, 1005–1013.
Wells, M. A. (1973) *Biochemistry 12*, 1086–1093.

Characterization of an Early Intermediate in the Folding of the α Subunit of Tryptophan Synthase by Hydrogen Exchange Measurement†

Anne M. Beasty and C. Robert Matthews*
Department of Chemistry, The Pennsylvania State University, University Park, Pennsylvania 16802
Received October 26, 1984

ABSTRACT: The development of the hydrogen bonding network in the early stages of the folding of the α subunit of tryptophan synthase was monitored with a hydrogen exchange technique. The orders of magnitude difference between the rapid conversions of the unfolded forms to two stable intermediates (milliseconds) and the subsequent slow conversions of the intermediates to the native form (>100 s) was used to selectively label with tritium the hydrogen bonds that form in the first 30 s of folding at 0 °C. Rapid removal of the tritiated solvent by gel filtration ensured that hydrogen bonds formed in subsequent folding reactions would be unlabeled. Limited proteolysis and separation of peptides by high-pressure liquid chromatography permitted the determination of the amount of label retained in individual peptides by scintillation counting. Peptides 1–70 and 71–188, which when covalently linked comprise the stable amino domain in the native conformation, retain 91% and 93%, respectively, of the label retained when the protein is allowed to completely refold in tritiated solvent. Peptide 189–268, the marginally stable carboxyl domain, only retains 43% of the label. The striking difference in retention of label confirms the independent folding of these two domains and shows that the kinetic intermediates that appear in the folding of α subunit correspond to structural domains in the native conformation. The near-equality of the labeling of the two peptides comprising the amino domain shows that this domain folds as a single entity and that subdomain folding is unlikely. Refolding into consecutively higher final urea concentrations shows that the amino domain consistently retains a greater fraction of label than does the carboxyl domain. This result demonstrates the greater stability of the amino domain.

The mechanism by which the amino acid sequence of a protein directs the folding to the native functional conformation is not known in detail for any protein. This lack of understanding is due in part to the high cooperativity of the unfolding transition and the concomitant absence of stable intermediates and in part to the rapid rate of folding which precludes the use of high-resolution techniques such as X-ray or nuclear magnetic resonance (NMR)[1] spectroscopy. The proliferation of DNA sequences for structural genes makes clear the desirability of being able to predict the three-dimensional conformation of a protein from its primary sequence.

In an effort to elucidate the mechanism of folding, we have been investigating the effects of single amino acid replacements on the folding and stability of the α subunit of tryptophan synthase from *Escherichia coli*. Over two dozen point mutations have been isolated and identified by Yanofsky and his colleagues (Yanofsky et al., 1967; Yanofsky & Horn, 1972; Murgola & Yanofsky, 1974). The basic premise of our work is that amino acids that play key roles in folding and stability will have observable effects on the kinetic and equilibrium properties of the unfolding reaction in vitro. Obviously, the correct interpretation of such effects depends upon a folding model for the wild-type protein that is as detailed as possible in both kinetic and structural terms.

Previous studies on the folding of the α subunit have proposed that one or more stable intermediates appear during folding (Yutani et al., 1979, 1982; Matthews & Crisanti, 1981; Crisanti & Matthews, 1981; Matthews et al., 1983). The structural interpretation of the principal intermediate is based upon the observation of Higgins et al. (1979) that the protein can be converted to two fragments by limited tryptic digestion. The amino domain consists of residues 1–188 and the carboxyl domain of residues 189–268. The isolated fragments spontaneously complement to form active enzyme. The principal intermediate has been proposed to consist of a folded amino domain and an unfolded carboxyl domain on the basis of both difference UV and circular dichroism spectroscopies. Supporting this hypothesis are the recent results of Miles et al. (1982) on the stabilities of the isolated domains to guanidine hydrochloride unfolding. The amino domain was shown to be substantially more stable than the carboxyl domain. Also,

† This work was supported by U.S. Public Health Service Grant GM 23303. C.R.M. is the recipient of Research Career Development Award K04 AG00153 from the U.S. Public Health Service.

[1] Abbreviations: NMR, nuclear magnetic resonance; $NaDodSO_4$, sodium dodecyl sulfate; HPLC, high-pressure liquid chromatography; $(NH_4)_2SO_4$, ammonium sulfate; DTE, dithioerythritol; Na_2EDTA, disodium ethylenediaminetetraacetate; NaOAc, sodium acetate; $^3H_{rem}$, moles of tritium remaining per mole of protein or peptide; Gdn-HCl, guanidine hydrochloride.

0006-2960/85/0424-3547$01.50/0

the unfolding transition of the isolated amino domain is coincident with the second of the two transitions that are observed in the intact protein. Thus, there appears to be a good correlation between the structure of the principal stable intermediate and the state of folding of the structural domains in the native conformation.

An important question for the folding mechanism of the α subunit is the relationship between the intermediates generated in the kinetic refolding experiment and the intermediates that are present at equilibrium. Our approach toward answering this question has been to take advantage of the fact that the exchange of protons involved in hydrogen bonds is slowed by orders of magnitude compared to solvent-exposed protons (Hvidt & Nielsen, 1966). Thus, by refolding in the presence of tritiated water, performing a proteolytic degradation, and rapidly separating the fragments, one might expect to determine the location and extent of hydrogen bonding in these kinetically derived intermediates. Comparison of the results with those from the completely labeled native conformation provides a means of determining the relationship between these structures.

Materials and Methods

Enzyme Purification and Characterization. The wild-type α subunit of tryptophan synthase from *E. coli* (EC 4.2.1.20) was isolated from strain B8/F'B8 containing plasmid pBN55 according to the methods of Kirschner et al. (1975) with modifications (Matthews et al., 1983).

The purity of the protein was ascertained by both Na-DodSO$_4$–polyacrylamide and native polyacrylamide gel electrophoreses; the electrophoretograms exhibited one symmetrical band indicative of only a single protein component. The purified enzyme was stored as a suspension containing 70% saturated $(NH_4)_2SO_4$ at 4 °C. The activity of the α subunit was measured by its ability to enhance the activity of the β_2 subunit in the condensation of indole and serine to form tryptophan; the reported maximum specific activity of the α subunit in this assay is 5500 units/mg (Kirschner & Wiskocil, 1975). The specific activities of preparations of the wild-type enzyme used in this study were 4500 ± 500 units/mg. Protein concentration was determined from the optical spectrum by using a specific absorption coefficient $\epsilon^{1\%}_{278\,nm} = 4.4$ (Adachi et al., 1974). The molar extinction coefficient was calculated to be 12 600 M^{-1} cm^{-1} by using a molecular weight of 28 700 (Yanofsky et al., 1967; Li & Yanofsky, 1972).

Chemicals and Reagents. Tritiated water of specific activity 0.5 Ci mL^{-1} and Aquasol-2 were purchased from New England Nuclear. Sephadex G-25-80 was purchased from Sigma Chemical Co. Acetonitrile was purchased from J. T. Baker and was HPLC grade. Trifluoroacetic acid (sequanal grade) and constant-boiling HCl were purchased from Pierce Chemical Co. Ultrapure urea was purchased from Schwarz/Mann and was used without further purification. All other chemicals were purchased from commercial sources and were reagent grade. All solutions were prepared by using twice-distilled deionized water.

Hydrogen Exchange of Native α Subunit. To produce α subunit which is completely tritiated at all exchangeable proton positions, the protein was incubated with tritiated water at 20.0 mCi mL^{-1} in 1.5 M urea, 10.0 mM potassium phosphate, 1.0 mM Na_2EDTA, and 1.0 mM DTE, pH 7.8 and 25 °C, for at least 10 h. Since at this concentration of urea the α subunit is only marginally stable (Matthews & Crisanti, 1981), the rate of introduction of tritium into the exchangeable positions is facilitated. The urea concentration was then lowered to 0.5 M urea by diluting this solution with an appropriate volume of 10.0 mM potassium phosphate, 1.0 mM Na_2EDTA, and 1.0 mM DTE, pH 7.8, 25 °C, with tritiated water at 20.0 mCi mL^{-1} for 1.0 h. At 0.5 M urea, where the α subunit is substantially more stable, the exchange of hydrogens involved in secondary and tertiary structure is substantially retarded.

The hydrogen exchange techniques used were those developed by Englander (Englander & Englander, 1972). Tritiated α subunit was separated from tritiated solvent by using gel filtration in jacketed columns (1.0 × 8.0 cm). Exchange data at times less than 20 min were obtained with a single column; data at longer times used a two-column technique. Effluent samples were collected through the protein peak. Samples ranged in size from two to four drops and were diluted to 1.0 mL for analysis. The protein concentration in each sample was calculated from the optical spectrum and the known extinction coefficient. Tritium levels were determined by counting 0.2 mL in 10.0 mL of Aquasol-2 in a Beckman Model LS7500 liquid scintillation counter. These data, together with a determination of the tritium level in the initial equilibration mixture, allow the calculation of hydrogen atoms remaining per molecule by the equation (Englander & Englander, 1972):

$$^3H_{rem} = \frac{111}{1.21}\frac{C}{C_0}\frac{\epsilon}{OD}$$

where 111 is the gram-atom concentration of hydrogen in H_2O and C_0 is the tritium level in the exchange-in mixture in counts per minute. C is the tritium level in the sample in counts per minute, ϵ is the molar extinction coefficient, OD is the absorbance of the α subunit at 278 nm, and the factor of 1.21 corrects for the tritium–hydrogen equilibrium isotope effect on the peptide NH group (Englander & Poulsen, 1969).

Tritium Labeling of Intermediates. Lyophilized α subunit was dissolved to a final concentration of 60.0 mg mL^{-1} in a solution containing 6.0 M urea, 10.0 mM potassium phosphate, 1.0 mM Na_2EDTA, 1.0 mM DTE, pH 7.8, and tritiated water at 20.0 mCimL^{-1}. The sample was allowed to incubate at 25 °C for 1 h to affect complete exchange of all peptide NH protons. An aliquot of this urea solution was cooled to 0 °C for 5 min prior to refolding. Refolding was initiated by a 10-fold dilution into buffered solvent at 0 °C containing 10.0 mM potassium phosphate, 1.0 mM Na_2EDTA, 1 mM DTE, pH 7.8, and tritiated water at 20 mCi mL^{-1}, and no urea. To remove unbound tritium and initiate exchange-out, 0.5 mL of this dilution mixture was immediately run through a gel filtration column at 0 °C. Elution of the protein from the column typically required about 30 s; the protein is removed from the tritiated solvent in the first few seconds of this process. The running buffer contained 50.0 mM NaOAc and 0.2 mM DTE, pH 5.5. The protein pooled from this column was then allowed to exchange-out for 6 h at pH 5.5 and 0 °C prior to digestion and peptide separation. This incubation allowed the α subunit to fold completely. The refolding rate constants for the two slow steps in folding do not depend on pH in the range from pH 5 to pH 8 at urea concentrations less than 1 M urea (Crisanti & Matthews, 1981). From the rate constant at 25 °C and the previously measured activation energies (Crisanti & Matthews, 1981), the relaxation times at 0 °C can be calculated to be 620 and 3450 s. Thus, the 6-h incubation represents nine half-lives for the slower of the two phases in folding. Also, the 6-h exchange-out exposes the most stable group of hydrogen bonded amide protons which the experiment is designed to monitor.

To compare the relative stabilities of the amino and carboxyl domains in the intermediate species, the experimental design was altered. As before, the α subunit was initially unfolded

by urea, and the NH positions were labeled with tritium. Refolding to different final urea concentrations was initiated by calculated dilutions into buffered solvent containing 10 mM potassium phosphate, 1.0 mM Na_2EDTA, 1.0 mM DTE, pH 7.80 at 0 °C, and tritiated water at 20 mCi mL^{-1}. To remove unbound tritium and initiate exchange-out, 0.5 mL of this dilution mixture was immediately run through a gel filtration column at 0 °C. The running buffer contained 50.0 mM NaOAc, 0.2 mM DTE, pH 5.5, and the same concentration of urea as the dilution mixture. The necessity for maintaining the final urea concentration while the tritium was removed is that the stability of the intermediates and their ability to retain tritium in hydrogen bonds depend on the final urea concentration (see Results). Thus, it is important to keep the urea concentration constant until the tritium has been removed from the solvent.

The protein collected from this column was allowed to incubate at pH 5.5 and 0 °C for a total of 4 min, including the time on the gel filtration column. The purpose of this incubation was to exchange-out all protons not involved in intraprotein hydrogen bonds. The half-time for exchange of the amide protons in poly(DL-alanine) is 38 s at pH 5.5 and 0 °C (Englander & Poulsen, 1969). The sample was then applied to a second gel filtration column run with 50.0 mM NaOAc, 0.2 mM DTE, pH 5.5, and 0.0 M urea at 0 °C. This step removes the urea and permits all protein samples to fold under the same final conditions. The protein from this column was then allowed to exchange-out for 6 h at pH 5.5 and 0 °C prior to digestion and peptide separation.

Complete Tritium Labeling of the α Subunit. An aliquot of the above solution of unfolded α subunit in 6 M urea was refolded in the presence of tritium by diluting 1:10 with 10.0 mM potassium phosphate, 1.0 mM Na_2EDTA, 1.0 mM DTE, pH 7.6, and tritium at 20.0 mCi mL^{-1}. Complete refolding was achieved by incubating the sample for 60 min at 25 °C. Exchange-out at pH 5.5 and 0 °C was initiated on a gel filtration column as described above for native α subunit. The pooled protein was allowed to exchange-out for 6 h at pH 5.5 and 0 °C.

Digest Conditions. Partial tryptic digestion of the α subunit was achieved by adding trypsin at a 1:10 enzyme:substrate ratio (mg/mg) for 6.0 min at pH 5.5 and 0 °C. The digestion was quenched by addition of concentrated acetic acid (20% v/v).

HPLC Conditions. A 0.25-mL aliquot of the partially digested protein was loaded onto a 3.0 × 300 mm Bondapak C-18 HPLC column to separate the peptide fragments. The column was maintained at 12 ± 1 °C in a water bath. Peptides were eluted with the following gradient at a flow rate of 2 mL min^{-1}. Solvent A was 5.0% CH_3CN and 0.1% CF_3COOH in water. Solvent B was 0.08% CF_3COOH in acetonitrile.

time (min)	% solvent A	% solvent B
0	65	35
13	52	48
19	52	48
31	40	60

The eluted peptides were detected by absorbance at 220 nm. Peak areas were obtained by automated integration.

Amino Acid Analysis. Peptide samples were lyophilized in hydrolysis tubes and hydrolyzed in vacuo for 24 h at 110 °C in 1.0 mL of constant-boiling HCl and 1.0 mM phenol. The resulting amino acids were determined by a reverse-phase separation according to the procedure of Hill et al. (1979).

Determination of $^3H_{rem}$ per Peptide. The peak areas of the peptide fragments were correlated with molar amounts obtained from amino acid analysis. Peptide concentrations in each sample were then determined from their peak areas and respective volumes. Tritium levels in the peptide solutions were determined by counting 0.5 mL in 10.0 mL of Aquasol-2. These data, together with a count of the tritium level in the initial dilution mixture of α subunit, allow the calculation of hydrogen atoms still unexchanged per peptide, $^3H_{rem}$.

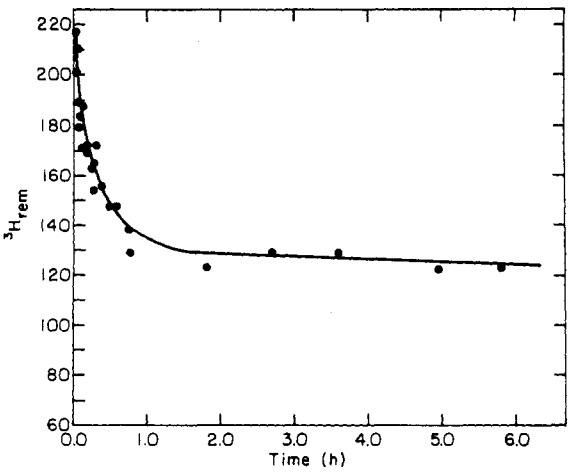

FIGURE 1: Exchange-out kinetics of native α subunit measured by $^3H_{rem}$ per molecule in 50 mM NaOAc and 0.2 mM DTE, pH 5.5, 0 °C.

RESULTS

The application of hydrogen exchange to determine the conformations of the stable intermediates in the folding of the α subunit of tryptophan synthase depends upon the ability to examine the behavior of a specific group of protons. The simplest procedure is to study the protons that are the slowest to exchange. For convenience, we will refer to this group of slowly exchanging protons as the core set.

Operationally, one identifies this set of protons by performing a classic exchange-out experiment on fully tritiated protein under specific conditions (Englander & Englander, 1972). Since the determination of the amount of the label retained in the amino and carboxyl domains requires limited proteolysis and separation of the domains, it is advantageous to choose conditions in which hydrogen exchange is minimized (Englander & Englander, 1972). In this way, one can maximize the retention of label at nitrogens whose hydrogen bonds will be broken during digestion or separation. Since the α subunit undergoes an acid-induced unfolding transition below pH 5 (Yutani et al., 1980), the exchange-out was performed at pH 5.5 and 0 °C.

The results of the exchange-out experiment for the fully tritiated α subunit are shown in Figure 1. At very short times, between 200 and 217 protons are protected. Hydrogen bonding is the major mechanism providing protection against exchange (Hvidt & Nielsen, 1966). Assuming that 80% of these protected protons are involved in secondary structure (Richardson, 1981), it appears that 160–170, approximately 60%, of the 268 amino acids participate in secondary structure. As exchange proceeds, the number of protected tritium atoms per molecule decreases significantly up to about 1 h at pH 5.5, 0 °C. After 1 h, further exchange proceeds very slowly. A core set of approximately 125 ± 4 stable protons is protected from exchange after 6 h. It is the behavior of this group of protons that will be studied in the folding reaction of the α subunit.

The proposed procedure requires proteolytic digestion of the protein and rapid separation of the fragments. Higgins et al.

Table I: Amino Acid Analysis of Three Major Fragments of a Limited Tryptic Digestion of the α Subunit[a,b]

	peptide I		peptide II		peptide III	
amino acid	observed	predicted for residues 188–268	observed	predicted for residues 71–188	observed	predicted for residues 1–70
Asp	5.2	6	11.7	10	6.3	6
Glu	8.4	8	11.8	11	10.0	10
Ser	3.9	4	4.4	4	3.4	3
His	1.9	2	1.8	2	0.0	0
Ala	17.0	17	16.0	16	7.0	7
Arg	1.1	1	6.8	7	3.4	3
Tyr	1.2	1	4.9	5	1.2	1
Val	4.6	5	9.3	10	2.4	2
Met	2.0	2	1.8	2	0.9	1
Ile	5.6	7	5.8	7	4.8	6
Leu	6.8	7	10.6	11	8.8	9
Phe	1.7	2	5.4	6	3.2	4
Lys	6.3	7	2.5	3	2.7	3

[a] Pro, Cys, Trp, Thr, and Gly were not determined. [b] All observed values were normalized to alanine.

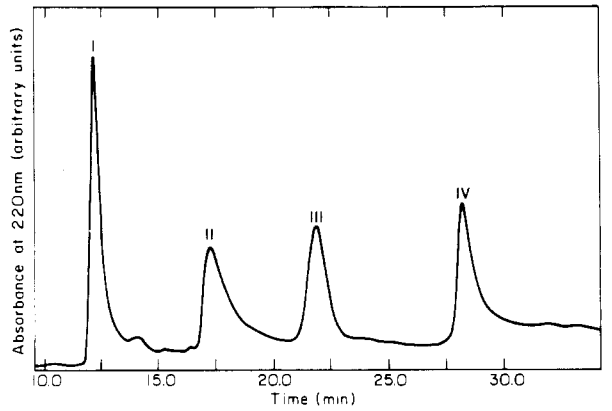

FIGURE 2: HPLC chromatogram of a limited tryptic digestion of the α subunit of tryptophan synthase. The separation procedure is described under Materials and Methods.

(1979) have previously found that a limited tryptic digestion of the $\alpha_2\beta_2$ holoenzyme results in a single cleavage in the backbone of the α subunit at Arg-188. To minimize subsequent separation problems in our experiments, it was advantageous to omit the β_2 subunit during digestion. Therefore, it was necessary to determine the limited tryptic digestion pattern of the isolated α subunit.

The high-pressure liquid chromatography chromatogram of the products of a limited tryptic digestion of the α subunit of pH 5.5, 0 °C, is shown in Figure 2. The identity of the peaks eluting at 12.5, 17.5, and 22 min, designated I, II, and III in Figure 2, was determined by amino acid composition. The results are shown in Table I.

The amino acid composition of the peptide comprising peak I agrees very well with that expected for peptide 189–268. This is the same cleavage obtained by Higgins et al. (1979) for the $\alpha_2\beta_2$ holoenzyme and demonstrates that the carboxyl domain can be prepared and cleanly isolated by digestion of the isolated α subunit. The amino acid composition of peak II agrees with that expected for peptide 71–188, and that for peak III agrees with that expected for peptide 1–70. In control experiments, undigested α subunit elutes at 28 min under these conditions and corresponds to peak IV. NaDoDSO$_4$–polyacrylamide gels of the products of this limited tryptic digestion show bands at the positions expected for cleavages at Arg-70 and Arg-188 (N. Tweedy, unpublished results). Miles (E. Miles, personal communication) has recently found that trypsin can cause partial cleavage at Arg-70 in the $\alpha_2\beta_2$ holoenzyme. Thus, the limited tryptic digestion of the isolated α subunit yields the intact carboxyl domain obtained previously and two fragments which comprise the amino domain as defined by Higgins et al. (1979). The division of the amino domain into two peptides will prove to be very useful in understanding the behavior of this domain during folding (see below).

Table II: $^3H_{rem}$/Peptide for Labeled α Subunit

	$^3H_{rem}$/peptide[a,b] for residues			
method	1–70	71–188	189–268	total $^3H_{rem}$
intermediates labeled	19.0 ± 3.7	75.0 ± 4.0	10.5 ± 1.4	99.2 ± 5.6
fully labeled	20.8 ± 2.5	80.2 ± 7.5	24.1 ± 2.9	125.1 ± 8.4
ratio[c]	91 ± 21	93 ± 10	43 ± 8	79 ± 7

[a] Following labeling, the protein was exchanged-out for 6 h at pH 5.5 and 0 °C. [b] Errors refer to standard deviations. [c] Ratio refers to intermediates labeled/fully labeled. Values are given in percent.

The distribution of the core set of 125 highly protected protons in the 3 tryptic peptides was then determined for a sample of α subunit that had been fully tritiated and exchanged-out for 6 h at pH 5.5, 0 °C. The results are shown in Table II. Peptide 1–70 contained 20.8 ± 2.5 such protons per molecule, peptide 71–188 contained 80.2 ± 7.5 protons, and peptide 189–268 contained 24.1 ± 2.9 protons. The total of 125.1 ± 8.4 protons is in excellent agreement with the value obtained for the intact protein, 125 ± 4 protons, showing that there is no measurable loss of trapped tritium during digestion and separation. This is surprising since the half-time for exchange of an exposed NH proton at pH 5.5, 0 °C, is approximately 40 s. Apparently, the two tryptic cleavages do not cause the α subunit to unfold and expose the core set of protons to solvent. The retention of protected tritium during separation on HPLC is very likely due to a combination of the effects of low temperature (12 °C), acidic pH, and the presence of an organic cosolvent (acetonitrile).

The application of hydrogen exchange to study the conformations of the intermediates that appear in the folding of the α subunit depends upon the kinetic properties of the previously proposed folding model (Matthews et al., 1983) (Figure 3). In this model, U_1 and U_2 are two unfolded forms, I_1, I_2, and I_3 are stable intermediates, and N is the native form. From hydrodynamic and difference UV studies (Matthews et al., 1981), it has been proposed that I_1 has a compact but not nativelike folded conformation and that I_2 has a folded amino domain and an unfolded carboxyl domain in the transition zone. I_3 differs from I_2 only in the state of isomerization of a proline residue, probably in the carboxyl domain. The relative U_1:U_2 ratio is 1:4. The unfolded forms U_1 and U_2 are thought to differ in the state of isomerization of the peptide bond at one or more X–Pro linkages. The measured or estimated time constants for each transition at 25 °C are in-

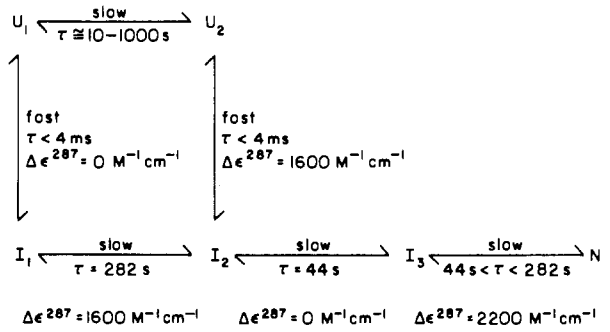

FIGURE 3: Proposed folding model for the α subunit of Trp synthase. The relaxation times measured or estimated at 25 °C, pH 7.8, for each of the transitions are shown. The calculated changes in the extinction coefficient at 287 nm in the folding direction for each of the transitions are also shown (Matthews & Crisanti, 1981).

dicated as are the expected values for $\Delta\epsilon_{287}$. The time constants for the $I_1 \rightarrow I_2$, $I_2 \rightarrow I_3$, and $I_3 \rightarrow N$ transitions have been measured directly. The range of values for the $I_3 \rightarrow N$ transition reflects the dependence of this parameter on the final urea concentration. Below 1.5 M urea, the refolding of I_2 is rate limited by the urea-independent conversion of I_2 to I_3; the conversion of I_3 to N presumably continues to increase in rate below 1.5 M urea but cannot be observed directly. The rapid collapse at U_1 to I_1 and of U_2 to I_2 has been deduced from the observation that the transition between the intermediates to the unfolded forms is fast on the NMR time scale (C. L. Froebe and C. R. Matthews, unpublished results). The relaxation time for the U_1 to U_2 conversion has not been measured; the range of values shown is that expected for proline isomerization at 25 °C (Grathwohl & Wuthrich, 1981).

Given this kinetic scheme, it is straightforward to design an experiment that reveals the fraction of the core set of protons that are trapped in the rapid collapse of U_1 and U_2 to I_1 and I_2. The protein is initially unfolded in a buffered solvent containing 6.0 M urea and 20 mCi mL^{-1} ^3H_2O; all the NH positions are fully equilibrated with tritium. Refolding is initiated by a 10-fold dilution into buffered solvent at 0 °C containing no urea and 20 mCi mL^{-1} ^3H_2O. After the solution is mixed, the protein is loaded within 30 s onto a gel filtration column preequilibrated at pH 5.5 and 0 °C. Excess tritium is removed by passage through this column. The protein is then allowed to refold at pH 5.5 for 6 h at 0 °C. This step serves the dual purpose of allowing the α subunit to completely refold to native subunit and of exchanging-out all protected tritium with the exception of the core set of protons. From the activation energies of the two slow steps in folding (Matthews et al., 1983), one can estimate that the time constants at 0 °C will be 3450 and 620 s for the I_1 to I_2 and the I_2 to I_3 transitions, respectively. These rates have been previously observed to be independent of pH below neutral pH (Crisanti & Matthews, 1981). Thus, during the 30-s exposure to tritium after refolding is initiated, only the protons that are protected by hydrogen bonds or by similar means during conversion of U_1 to I_1 and of U_2 to I_2 will be labeled with tritium.

The refolded protein was then subjected to limited tryptic digestion, and the fragments were separated on HPLC. The tritium content of each of the peptides was then determined by scintillation counting. The results are shown in Table II. There is evidently a significant difference in the behavior of the amino and carboxyl domains of the α subunit. Peptide 1–70 contains 91% of the core set of protons, and peptide 71–188 contains 93% of this set. Considering the error limits, one can conclude that almost the entire core set of protons

Table III: $^3H_{rem}$/Peptide for Labeled Folding Intermediates in the α Subunit at Different Final Urea Concentrations

	$^3H_{rem}$/peptide[a,b] for residues		
[urea] (M)	1–70	71–188	189–268
0.45	17.8 ± 2.3	ND[c]	9.5 ± 1.5
0.75	16.8 ± 0.8	ND[c]	6.8 ± 1.4
1.00	16.0 ± 0.5	63.0 ± 8.1	5.2 ± 0.6
1.25	11.3 ± 0.3	47.3 ± 1.6	2.9 ± 0.2
1.50	12.2 ± 0.6	48.2 ± 2.3	2.8 ± 0.2

[a] Following labeling, the protein was exchanged-out for 6 h at pH 5.5 and 0 °C. [b] Errors refer to standard deviations. [c] Not determined.

forms in the amino domain within 30 s. The observation that the two amino domain fragments have proportionately the same tritium contents within experimental error shows that the sequence from 1 to 188 acts as a single folding unit in this experiment. In contrast, the carboxyl domain retains only 43% of the core set of protons under these conditions.

The tritium labeling pattern supports the hypothesis that the amino and carboxyl domains fold independently. It also demonstrates that the folding intermediates that appear within 30 s reflect the domain structure of the native conformation.

Although the carboxyl domain retains a markedly lower fraction of its core set of protons than the amino domain, the observed value of 43% shows that substantial hydrogen bonding occurs at 0 °C and at low urea concentration. Thus, it appears that this domain either is partially folded or is a mixture of folded and unfolded forms under these conditions. Another possibility is that the carboxyl domain folds more slowly and hydrogen exchange can compete more effectively.

This result prompted an experiment to test the relative stabilities of the amino and carboxyl domains. As before, the α subunit was initially unfolded by urea, and the NH positions were labeled with tritium. The refolding reaction was again initiated by dilution; however, the buffer into which the protein was diluted contained sufficient urea so as to vary the final urea concentration from 0.45 to 1.50 M. The sample was thoroughly mixed and loaded onto a gel filtration column preequilibrated at 0 °C with a solvent buffered at pH 5.5 and containing the same urea concentration. This column removes the excess tritium but maintains the urea concentration and, therefore, the conformation of the protein. This consideration is important since Miles et al. (1982) have concluded that the isolated carboxyl domain is not completely folded at 25 °C, pH 7.8, in the absence of denaturant. Thus, it is expected that the conformation of this domain will be sensitive to the urea concentration, even at low concentrations. The pooled fractions of protein were incubated at 0 °C for 4 min (from the start of the gel filtration column) to allow exchange of any solvent-exposed NH protons. The sample was then passed through a second gel filtration column equilibrated with buffer at pH 5.5, 0 °C, and containing no urea. The sample was allowed to refold and exchange-out for 6 h at pH 5.5, 0 °C, in order to expose the core set of protons.

The tritium remaining per molecule for each of the three tryptic peptides at a series of final urea concentrations is shown in Table III. As the final urea concentration is increased from 0.45 to 1.50 M, each of the peptides shows a decreased amount of tritium trapped in the core set of protons. However, the relative decrease for the carboxyl domain is much greater than that for either of the two amino domain fragments. This effect can be graphically demonstrated by normalizing the values relative to the values for the fully labeled core set of protons (Table II). The result is shown in Figure 4.

The following observations can be made: (1) The amino domain forms a significantly higher percentage of its most

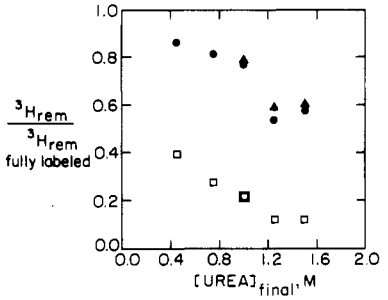

FIGURE 4: Dependence of the fraction of tritium remaining for each of the three tryptic peptides as a fraction of the final urea concentration for residues 1–70 (●), 71–188 (▲), and 189–268 (□). Values have been normalized relative to the values for the fully labeled peptides.

stable hydrogen bonds than the carboxyl domain at the same final urea concentration. This result is consistent with the hypothesis that the amino domain is more stable. Since the dependence of the tritium remaining in the fragments on final urea concentration is too complex to be fit by a simple model, the extrapolation of the data to 0 M urea cannot be done. However, it is clear that even at 0.45 M urea the core set of protons in the amino domain is formed almost entirely within 30 s. (2) At 0.45 M urea, nearly 40% of the core set of protons in the carboxyl domain are protected from exchange. Apparently, at sufficiently low urea concentrations, substantial secondary structure forms in the carboxyl domain. (3) As the final urea concentration is increased, the tritium trapped in stable hydrogen bonds decreases monotonically in both domains. The greater stability of the amino domain is evident in the fact that it retains 60% of its core set of protons at 1.5 M urea while the carboxyl domain retains only 12%. It appears that the carboxyl domain is almost completely unfolded at 1.5 M urea. (4) As discussed for the previous experiment, the close agreement between the relative amount of tritium retained by both the fragments that comprise the amino domain clearly shows that this domain acts as a single functional unit in the folding mechanism.

It should be noted that urea can affect hydrogen exchange rates of amide hydrogens. Measurements on poly(DL-alanine) show that the rate is slowed by ~15% for urea concentrations less than 1.5 M (D. Loftus, P. Kim, and R. Baldwin, unpublished results). Correction of the data for an effect of this magnitude would not alter the above conclusions.

Discussion

An important question regarding the mechanism of protein folding is the relationship between the conformation of the intermediates that appear in refolding and the native conformation. Creighton has found obligatory nonnative disulfide-bonded intermediates in the refolding of reduced, unfolded bovine trypsin inhibitor (Creighton, 1977). Thus, the refolding reaction does not have to be a strict progression from the unfolded to the native conformation.

The data presented in this paper on the folding of the α subunit of tryptophan synthase show that the amino and carboxyl domains independently establish their hydrogen bonding networks and, presumably, their secondary structures. The simplest explanation for this result is that the two domains fold independently of each other. This conclusion is supported by the observation that refolding in the presence of increasing concentrations of denaturant results in a preferential loss of tritium label in the carboxyl domain. If the α subunit were to fold as a single unit, uniform loss of label throughout the entire polypeptide would be expected. The selective loss of tritium in the carboxyl domain is consistent with the hypothesis that the carboxyl domain is substantially less stable than the amino domain. Therefore, in the unfolding transition zone, the kinetic intermediate would be expected to resemble the principal equilibrium intermediate which has a folded amino domain and an unfolded carboxyl domain. Taken together, these observations suggest that the folding proceeds in a progressive fashion through intermediates whose conformations and relative stabilities are very similar to those for domains found in the native conformation.

Since a kinetic scheme in which one or more rapid steps precede the slowest step in folding allows the intermediates to fully equilibrate, it is very likely that the conformations that the amino and carboxyl domains ultimately adopt in the native form must also be the most stable conformations for those segments of the polypeptide chain. This observation provides further experimental support for the hypothesis that the native conformation corresponds to the global free energy minimum (Anfinsen, 1973) and increases confidence in computational efforts to determine the native conformation of a given amino acid sequence by energy minimization.

It is also interesting that although the isomeric configuration of X–Pro peptide bonds can serve to limit the rate of folding in the α subunit (Crisanti & Matthews, 1981), they do not appear to drastically alter the secondary structure that appears rapidly in the amino domain. Assuming that the cis:trans ratio at X–Pro peptide bonds is 1:4 and that all 13 such peptide bonds in the amino domain are trans in the native conformation, only 5.4% of the unfolded molecules have the correct isomeric configurations in the amino domain. Since nearly all the stable hydrogen bonds in the amino domain form rapidly at low urea concentration, it is evident that substantial hydrogen bonding occurs in species that contain incorrect isomers at X–Pro peptide bonds. This observation may be understood in part by noting that Pro residues often occur in turns where flexibility and exposure to solvent may tend to minimize the effect of an incorrect isomeric form on the secondary structure.

Although nonnative X–Pro peptide bonds do not appear to drastically alter the most stable elements of the secondary structure of the amino domain, they may well affect the tertiary structure and, thereby, decrease the stability. Miles et al. (1982) have found that the unfolding reaction occurring at higher Gdn·HCl concentration coincides with that of the isolated amino domain. For urea-induced unfolding, this transition begins around 3 M urea at 25 °C and has a midpoint near 4.5 M. Thus, one would have expected that this domain would protect its core set of protons at urea concentrations less than 1.5 M. The observation that only 60% of the core set remains at 1.5 M urea suggests that the kinetically derived domain is significantly less stable than its ultimate equilibrium conformation. Previous studies on the refolding of ribonuclease A have identified a native like intermediate, I_N, that has enzymatic activity but which does not have all of the prolines in their correct isomeric forms (Cook et al., 1979; Schmid, 1981; Schmid & Blascheck, 1981). This species unfolds more rapidly in Gdn·HCl-induced unfolding than does the native conformation (Schmid, 1983).

The kinetic model derived from optical experiments proposes that two intermediates, I_1 and I_2, appear in refolding. The experiments described in the present study do not permit an independent analysis of the hydrogen bonding networks in each intermediate. However, the significant amount of tritium trapped in the folded carboxyl domain at low urea refines our view on the conformation of this domain. Apparently, sub-

stantial secondary structure can form under conditions appropriate for the native form.

The significantly greater amount of label retained in the amino domain fragment 71–188, 80 $^3H_{rem}$/peptide, compared to fragment 1–70, 21 $^3H_{rem}$/peptide, shows that fragment 71–188 is richer in stable hydrogen bonds and presumably secondary structure than is fragment 1–70. Although a detailed analysis of this observation must await the crystallographic structure, application of the Chou–Fasman procedure (Chou & Fasman, 1978) for predicting secondary structure supports this hypothesis. Fragment 1–70 is predicted to have 36 residues in α helices, 9 in β sheets, and 12 in turns while fragment 71–188 is predicted to have 53 residues in α helices, 30 in β sheets, and 24 in turns (M. Hurle, unpublished results). The additional 50 residues in fragment 71–188 predicted to be involved with secondary structure may account in part for the difference in stable hydrogen bond content between these two fragments.

The hydrogen exchange method for monitoring protein folding has certain advantages and disadvantages compared to other methods commonly used. The biggest advantage is the potential for each amino acid, through its amide proton, to act as a label to follow the development of organized structure during folding. Optical techniques such as difference ultraviolet and fluorescence spectroscopies can only detect chromophores such as tyrosine and tryptophan and thus provide a more limited view of the folding process. Another potential advantage for hydrogen exchange is the possibility that the labeled protein can be cleaved into fragments and the tritium content of the individual fragments determined. Thus, it is in principle possible to obtain information on the hydrogen bonding pattern and, thereby, the conformation of intermediates in folding. This is possible only in a very limited way with difference ultraviolet and fluorescence spectroscopies. Another advantage is the greater sensitivity of scintillation counting compared to the optical methods mentioned above. The corresponding reduction in the amount of protein required to perform a folding study could be an advantage in certain systems. Probably the biggest disadvantage is the inability to follow the folding process in real time. The necessity of doing stopped-time assays makes the collection of data laborious.

The folding of a number of other globular proteins has been proposed to proceed by the folding of structural domains (Högberg-Raibaud & Goldberg, 1977; Carrey & Pain, 1978; Johanson et al., 1981; Dautry-Versat & Garel, 1981; Zetina & Goldberg, 1982). Hydrogen exchange methods, which have previously proved to be of value in demonstrating the presence of transient intermediates in the folding of a single-domain protein, ribonuclease A (Schmid & Baldwin, 1979; Kim & Baldwin, 1980), are also a useful way of monitoring domain folding in systems with appropriate kinetic properties.

ACKNOWLEDGMENTS

We thank Joan Englander and Dr. S. W. Englander for their assistance with the hydrogen exchange procedures, Dr. Robert Baldwin and Peter Kim for their helpful comments, Thomas Stackhouse for preparation of purified α subunit, Neil Tweedy for assistance with an electrophoresis experiment, and Gail Wasson for typing the manuscript.

Registry No. Hydrogen, 1333-74-0; tryptophan synthase, 9014-52-2.

REFERENCES

Adachi, O., Kohn, L. D.; & Miles, E. W. (1974) *J. Biol. Chem. 249*, 7756.
Anfinsen, C. B. (1973) *Science (Washington, D.C.) 181*, 223.
Carrey, E. A., & Pain, R. M. (1978) *Biochim. Biophys. Acta 533*, 12.
Chou, P. Y., & Fasman, G. D. (1978) *Adv. Enzymol. Relat. Areas Mol. Biol. 47*, 45.
Cook, K. H., Schmid, F. X., & Baldwin, R. L. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 6157.
Creighton, T. E. (1977) *Prog. Biophys. Mol. Biol. 33*, 231.
Crisanti, M. M., & Matthews, C. R. (1981) *Biochemistry 20*, 2700.
Dautry-Versat, A., & Garel, J.-R. (1981) *Biochemistry 20*, 1396.
Englander, S. W., & Poulsen, A. (1969) *Biopolymers 7*, 379.
Englander, S. W., & Englander, J. J. (1972) *Methods Enzymol. 26*, 406.
Grathwohl, C., & Wuthrich, K. (1981) *Biopolymers 20*, 2623.
Higgins, W., Fairwell, T., & Miles, E. W. (1979) *Biochemistry 18*, 4827.
Hill, D. W., Walters, F. H., Wilson, T. D., & Stuart, J. D. (1979) *Anal. Chem. 51*, 1338.
Högberg-Raibaud, A., & Goldberg, M. E. (1977) *Biochemistry 16*, 4014.
Hvidt, A., & Nielsen, O. (1966) *Adv. Protein Chem. 21*, 287.
Johanson, K. O., Wetlaufer, D. B., Reed, R. G., & Peters, T., Jr. (1981) *J. Biol. Chem. 256*, 445.
Kim, P. S., & Baldwin, R. L. (1980) *Biochemistry 19*, 6124.
Kirschner, K., Wiskocil, R. L., Loehn, M., & Rezeau, L. (1975) *Eur. J. Biochem. 60*, 513.
Li, S.-L., & Yanofsky, C. (1972) *J. Biol. Chem. 247*, 1031.
Matthews, C. R., & Crisanti, M. M. (1981) *Biochemistry 20*, 784.
Matthews, C. R., Crisanti, M. M., Manx, J. T., & Gepner, G. L. (1983) *Biochemistry 22*, 1445.
Miles, E. W., Yutani, K., & Ogasahara, K. (1982) *Biochemistry 21*, 2586.
Murgola, E. J., & Yanofsky, C. (1974) *J. Mol. Biol. 86*, 775.
Richardson, J. S. (1981) *Adv. Protein Chem. 34*, 167.
Schmid, F. X. (1981) *Eur. J. Biochem. 114*, 105.
Schmid, F. X. (1983) *Biochemistry 22*, 4690.
Schmid, F. X., & Baldwin, R. L. (1979) *J. Mol. Biol. 135*, 199.
Schmid, F. X., & Blaschek, H. (1981) *Eur. J. Biochem. 114*, 111.
Yanofsky, C. (1967) *Harvey Lect. 61*, 145.
Yanofsky, C., & Horn, V. (1972) *J. Biol. Chem. 247*, 4494.
Yanofsky, C., Drapeau, G. R., Guest, J. R., & Carlton, B. C. (1967) *Proc. Natl. Acad. Sci. U.S.A. 57*, 296.
Yutani, K., Ogasahara, K., Suzuki, M., & Sugino, Y. (1979) *J. Biochem. (Tokyo) 85*, 915.
Yutani, K., Ogasahara, K., & Sugino, Y. (1980) *J. Mol. Biol. 144*, 455.
Yutani, K., Ogasahara, K., Kimura, A., & Sugino, Y. (1982) *J. Mol. Biol. 160*, 387.
Zetina, C. R., & Goldberg, M. E. (1982) *J. Mol. Biol. 157*, 133.