

changes and degree of saturation is due to different spectral properties of the two quaternary states of Hb upon oxygenation. This difference may be attributed at a deeper level to different optical properties of the  $\alpha$  and  $\beta$  chains in the two possible quaternary T and R states. Deconvolution of the Hb binding properties observed in the critical isosbestic region will help to characterize the functional properties of the T and the R states at intermediate stages of ligation.

#### ACKNOWLEDGMENTS

We thank Bo Hedlund for his assistance in preparing the Hb samples.

#### REFERENCES

- Adair, G. S. (1925) *J. Biol. Chem.* 63, 529-545.  
 Adams, M. L., & Schuster, T. M. (1974) *Biochem. Biophys. Res. Commun.* 58, 525-531.  
 Di Cera, E., Doyle, M. L., Connelly, P. R., & Gill, S. J. (1987a) *Biochemistry* 26, 6494-6502.  
 Di Cera, E., Robert, C. H., & Gill, S. J. (1987b) *Biochemistry* 26, 4003-4008.  
 Dolman, D., & Gill, S. J. (1978) *Anal. Biochem.* 87, 127-134.  
 Doyle, M. L., Di Cera, E., Robert, C. H., & Gill, S. J. (1987) *J. Mol. Biol.* 196, 927-934.  
 Enoki, Y., & Tyuma, I. (1964) *Jpn. J. Physiol.* 14, 280-298.  
 Frazier, R. D. B., & Suzuki, E. (1973) in *Physical Principles and Techniques of Protein Chemistry*, Academic Press, New York.  
 Gibson, Q. H. (1959) *Prog. Biophys. Biophys. Chem.* 9, 1-53.  
 Gill, S. J., Di Cera, E., Doyle, M. L., Bishop, G. A., & Robert, C. H. (1987) *Biochemistry* 26, 3995-4002.  
 Hayashi, A., Suzuki, T., & Shin, M. (1973) *Biochim. Biophys. Acta* 310, 309-316.  
 Hofrichter, J., Sommer, J. H., Henry, E. R., & Eaton, W. A. (1983) *Proc. Natl. Acad. Sci. U.S.A.* 80, 2235-2239.  
 Imaizumi, K., Imai, K., & Tyuma, I. (1978) *J. Biochem. (Tokyo)* 83, 1707-1713.  
 Magar, M. E. (1972) in *Data Analysis in Biochemistry and Biophysics*, Academic Press, New York.  
 Mills, F. C., Johnson, M. L., & Ackers, G. K. (1976) *Biochemistry* 15, 5350-5362.  
 Nasuda-Kouyama, A., Tachibana, H., & Wada, A. (1983) *J. Mol. Biol.* 164, 451-476.  
 Perutz, M. F. (1970) *Nature (London)* 228, 726-739.  
 Rifkind, J., & Lumry, R. (1967) *Fed. Proc., Fed. Am. Soc. Exp. Biol.* 26, 2325.  
 Williams, R. C., & Tsay, K. (1973) *Anal. Biochem.* 54, 137-145.

## Folding of Homologous Proteins: Conservation of the Folding Mechanism of the $\alpha$ Subunit of Tryptophan Synthase from *Escherichia coli*, *Salmonella typhimurium*, and Five Interspecies Hybrids<sup>†</sup>

Thomas M. Stackhouse,<sup>‡§</sup> James J. Onuffer,<sup>‡</sup> C. Robert Matthews,<sup>\*‡</sup> Syed A. Ahmed,<sup>||,⊥</sup> and Edith W. Miles<sup>||</sup>  
 Department of Chemistry, The Pennsylvania State University, University Park, Pennsylvania 16802, and Laboratory of Biochemical Pharmacology, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20205

Received June 26, 1987; Revised Manuscript Received September 21, 1987

**ABSTRACT:** The equilibrium and kinetic properties for the urea-induced unfolding of the  $\alpha$  subunit of tryptophan synthase from *Escherichia coli*, *Salmonella typhimurium*, and five interspecies hybrids were compared to determine the role of protein folding in evolution. The parent proteins differ at 40 positions in the sequence of 268 amino acids, and the hybrids differ by up to 15 amino acids from the *Escherichia coli*  $\alpha$  subunit. The results show that all the proteins follow the same folding mechanism and are consistent with a previously proposed hypothesis [Hollecker, M., & Creighton, T. E. (1983) *J. Mol. Biol.* 168, 409; Krebs, H., Schmid, F. X., & Jaenicke, R. (1983) *J. Mol. Biol.* 169, 619] that the folding mechanisms are conserved in homologous proteins. Analysis of the kinetic data suggests that the 15 positions at which the parent proteins differ in the amino folding unit, residues 1-188, do not play a role in a rate-limiting step in folding that has been previously identified as the association of the amino and carboxyl folding units [Beasty, A. M., Hurle, M. R., Manz, J. T., Stackhouse, T. S., Onuffer, J. J., & Matthews, C. R. (1986) *Biochemistry* 25, 2965]. One or more of the 25 positions at which the parent proteins differ in the carboxyl folding unit, residues 189-268, do appear to play a role in this same rate-limiting step.

**C**omparative studies on homologous proteins have generally focused on the effect of amino acid changes on the structure,

stability, and function (Rossman & Argos, 1981; Bajaj & Blundell, 1984). The results highlight individual residues that play key roles in substrate binding and/or catalysis as well as the general principles that relate amino acid sequence to secondary and tertiary structure.

A closely related question concerns the effects of these amino acid replacements on the folding mechanism. Comparison of the refolding kinetics of ribonucleases from a group of ruminants showed that the same mechanism applied for all variants (Krebs et al., 1983; Lang et al., 1986). From similar studies on another group of ribonucleases, Schmid et al. (1986)

<sup>†</sup> This research was supported by USPHS Grant GM 23303 (C.R.M.) and by Research Career Development Award 1 KO4 AG00153 (C.R.M.).

<sup>‡</sup> The Pennsylvania State University.

<sup>§</sup> Present address: Department of Biochemistry, University of California, Davis, Davis, CA 95616.

<sup>||</sup> National Institutes of Health.

<sup>⊥</sup> Present address: Department of Biochemistry, Bowman Gray School of Medicine, Winston-Salem, NC 27107.

concluded that one of the slow-folding phases is due to cis-trans isomerization at a specific X-Pro peptide bond. Therefore, for this collection of ribonucleases, the in vitro folding pathway is preserved. Hollecker and Creighton (1983) also found that the pathways of unfolding and refolding of three homologous protease inhibitors are closely related, suggesting that in this system, as well, folding pathways are conserved. This concept is important in refining an understanding of the rate-limiting steps in folding because, if it is correct, comparisons of the sequences of homologous proteins will emphasize the residues that play essential roles in this complex process.

The  $\alpha$  subunit of tryptophan synthase is an excellent system to test further the generality of the hypothesis on the conservation of the folding mechanism during evolution:

(1) The genes for both the *Escherichia coli* and *Salmonella typhimurium*  $\alpha$  subunits have been cloned into high copy number plasmids, providing a ready source of protein.

(2) Five interspecies hybrids have been made by fusing the genes from the parent organisms; the fusion proteins are enzymatically active (Schneider et al., 1981).

(3) Extensive equilibrium studies on the folding of the  $\alpha$  subunit have shown that the urea-induced (Matthews & Crisanti, 1981) and guanidine hydrochloride (Yutani et al., 1979) induced unfolding follow a multistate mechanism,  $N \leftrightarrow I \leftrightarrow U$ . The stable intermediate, I, is thought to have residues 1-188 folded into a nativelike conformation and residues 189-268 in an unfolded conformation (Miles et al., 1982; Crisanti & Matthews, 1981). The hybrids involve fusions at various positions in the amino folding unit, providing an opportunity to investigate the effect of combining different segments from *Escherichia coli* (EC)<sup>1</sup> and ST  $\alpha$  subunits on its folding and stability.

(4) A preliminary report on the X-ray structure has appeared (Hyde et al., 1987), indicating that a structural basis for interpretation will become available.

In a previous study on the stabilities of the two parent proteins and one of these hybrids to both guanidine hydrochloride and thermal unfolding (Yutani et al., 1984), it was reported that the stability of the hybrid reflects the stability of the parent protein from which the polypeptide segment is derived. In the present paper, the study of the parent and hybrid  $\alpha$  subunits was significantly extended both to include the four additional hybrids constructed by Schneider et al. (1981) and to investigate the effects on the relaxation times for the rate-limiting steps in folding. The results show that, similar to the ribonucleases and protease inhibitors, the folding mechanism is conserved across this set of proteins. Furthermore, in contradiction to the conclusion reached in the study by Yutani et al. (1984), the fusion of the EC and ST  $\alpha$  subunits has nonadditive effects on the stability.

## EXPERIMENTAL PROCEDURES

**Protein Purification.** The  $\alpha$  subunit (EC 4.2.1.20) was purified from *Escherichia coli* containing plasmid pBN 55 (a gift from Brian Nichols) and the 6-34, 8-32, 12-28, 14-26, and 15-25 hybrid  $\alpha$ -subunit proteins from *Escherichia coli* containing the appropriate plasmid (Schneider et al., 1981). *Salmonella typhimurium* LT2 (strain TB1533 TrpR 782 LEDC 1682 Trp B+ A+), a gift from Ronald Bauerle, was the source of tryptophan synthase  $\alpha_2\beta_2$  complex which was

purified (Miles et al., 1987) and then separated into  $\alpha$  and  $\beta_2$  subunits as described for the *Escherichia coli*  $\alpha_2\beta_2$  complex (Higgins et al., 1979).

The procedure used in purification was similar to that which has been described previously (Matthews et al., 1983) with the exception that a linear 0-1.0 M NaCl gradient on the initial DE-52 column was used to elute the 15-25 hybrid protein and a linear 0-0.3 M NaCl gradient was used for the remaining hybrid proteins. The purification of the ST  $\alpha$  subunit was identical with that for the EC  $\alpha$  subunit.

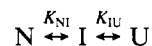
**Activity Assay.** The activities of all the  $\alpha$  subunits in this study were determined by measuring their ability to enhance the activity of the  $\beta_2$  subunit of tryptophan synthase from *Escherichia coli* in the condensation of indole and serine to form tryptophan; the maximum specific activity of EC  $\alpha$  subunit is 5500 units/mg in this assay (Kirschner et al., 1975). The activities of the EC, ST, 6-34, 8-32, 12-28, 14-26, and 15-25 hybrid  $\alpha$  subunits were 4500, 3400, 4400, 4000, 4800, 3000, and 3400 units/mg, respectively, with an error of  $\pm 10\%$ .

**Protein Purity and Concentration Determination.** The purity of each protein was shown by the appearance of a single band on Coomassie Blue stained native and sodium dodecyl sulfate-polyacrylamide gel electrophoresis. The protein concentration was determined by absorbance measurements at 278 nm using a molar extinction coefficient of  $12600 \text{ M}^{-1} \text{ cm}^{-1}$  (Matthews & Crisanti, 1981) for all the proteins in this study. The seven Tyr residues are strictly conserved as is the absence of Trp residues. The reversibility was determined by the recovery of the absorbance spectrum and always exceeded 90%. The protein concentration was held constant in any given experiment, generally in the range of 0.6-1.5 mg mL<sup>-1</sup>.

**Chemicals.** Ultrapure urea was purchased from Schwarz/Mann; all other chemicals were reagent grade. The buffer used in all folding experiments was 10 mM potassium phosphate, pH 7.8, 0.2 mM EDTA, and 1 mM 2-mercaptoethanol.

**Spectroscopic Methods.** Ultraviolet difference spectroscopy measurements were made at 25 °C on a Cary 118 CX spectrophotometer using the tandem cell technique (Herskovits, 1967). Equilibrium and kinetic methods for monitoring the unfolding and refolding reactions have been described previously (Crisanti & Matthews, 1981). Equilibrium measurements were made after all changes in absorbance had occurred, generally 1-2 h at 25 °C.

**Data Analysis.** The equilibrium unfolding results were fit to a three-state model:



where N, I, and U are native, intermediate, and unfolded forms, respectively,  $K_{NI} = [I]/[N]$ , and  $K_{IU} = [U]/[I]$ . The observed changes in extinction coefficient at 286 nm were converted to the apparent fraction unfolded,  $F_{app}$ , by  $F_{app} = (\epsilon_{obsd} - \epsilon_N)/(\epsilon_U - \epsilon_N)$  where  $\epsilon_{obsd}$  is the observed extinction coefficient at a given urea concentration and  $\epsilon_N$  and  $\epsilon_U$  are the extinction coefficients for the native and unfolded forms, respectively, at the same urea concentration.  $\epsilon_N$  and  $\epsilon_U$  showed small linear dependences on the urea concentration in their respective base-line regions. Linear extrapolations were made to obtain estimates for these values in the transition region. The apparent equilibrium constant for unfolding,  $K_{app}$ , was then calculated from  $K_{app} = F_{app}/(1 - F_{app})$ .  $K_{app}$  can also be defined in terms of the equilibrium constants from the three-state model:

$$K_{app} = \frac{K_{NI}K_{IU} + ZK_{NI}}{1 + (1 - Z)K_{NI}}$$

<sup>1</sup> Abbreviations: EC  $\alpha$  subunit,  $\alpha$  subunit of tryptophan synthase from *Escherichia coli*; ST  $\alpha$  subunit,  $\alpha$  subunit of tryptophan synthase from *Salmonella typhimurium*; EDTA, ethylenediaminetetraacetic acid.

Table I: Thermodynamic Parameters for the Urea-Induced Unfolding of *Escherichia coli*, *Salmonella typhimurium*, and Five Interspecies Hybrid  $\alpha$  Subunits

| $\alpha$ subunit | $\Delta G_{\text{NI}}^{\text{H}_2\text{O}}$<br>(kcal mol <sup>-1</sup> ) | $A_{\text{NI}}^b$ (kcal mol <sup>-1</sup> M <sup>-1</sup> ) | $\Delta G_{\text{IU}}^{\text{H}_2\text{O}}$<br>(kcal mol <sup>-1</sup> ) | $A_{\text{IU}}^b$ (kcal mol <sup>-1</sup> M <sup>-1</sup> ) | $C_{\text{mNI}}^c$ (M, urea) | $\Delta\Delta G_{\text{NI}}^d$<br>(kcal mol <sup>-1</sup> ) | $C_{\text{mIU}}^c$ (M, urea) | $\Delta\Delta G_{\text{IU}}^d$<br>(kcal mol <sup>-1</sup> ) |
|------------------|--|---|--|---|------------------------------|---|------------------------------|---|
| EC               | 5.7 ± 0.4  | 2.18 ± 0.15   | 5.0 ± 0.6  | 1.20 ± 0.15   | 2.62 ± 0.02                  |   | 4.18 ± 0.06                  |   |
| ST               | 6.1 ± 0.5  | 2.19 ± 0.16   | 6.2 ± 0.7  | 1.56 ± 0.16   | 2.79 ± 0.04                  | 0.4 ± 0.1   | 3.93 ± 0.05                  | -0.4 ± 0.1  |
| 6-34             | 5.4 ± 0.4  | 2.73 ± 0.23   | 4.8 ± 0.5  | 1.40 ± 0.14   | 1.98 ± 0.2                   | -1.8 ± 0.2  | 3.40 ± 0.04                  | -1.1 ± 0.1  |
| 8-32             | 5.4 ± 0.8  | 2.49 ± 0.39   | 6.2 ± 1.5  | 1.84 ± 0.43   | 2.17 ± 0.04                  | -1.1 ± 0.2  | 3.37 ± 0.07                  | -1.5 ± 0.3  |
| 12-28            | 6.4 ± 1.6  | 3.28 ± 0.84   | 4.0 ± 1.2  | 1.34 ± 0.39   | 1.95 ± 0.06                  | -2.2 ± 0.7  | 3.01 ± 0.12                  | -1.6 ± 0.4  |
| 14-26            | 5.0 ± 1.1  | 2.50 ± 0.57   | 6.3 ± 1.6  | 1.78 ± 0.45   | 2.01 ± 0.05                  | -1.5 ± 0.4  | 3.54 ± 0.08                  | -1.1 ± 0.3  |
| 15-25            | 3.9 ± 2.0  | 2.02 ± 1.05   | 3.7 ± 2.5  | 1.10 ± 0.71   | 1.92 ± 0.18                  | -1.4 ± 0.9  | 3.33 ± 0.29                  | -0.9 ± 0.6  |

<sup>a</sup>  $\Delta G_{\text{XY}}^{\text{H}_2\text{O}}$  is the free energy difference between conformations X and Y in the absence of urea; the units are kilocalories per mole. <sup>b</sup>  $A_{\text{XY}}$  is the urea dependence of the apparent free energy difference,  $\Delta G_{\text{XY}}^{\text{app}}$ , between conformations X and Y, defined by  $\Delta G_{\text{XY}}^{\text{app}} = \Delta G_{\text{XY}}^{\text{H}_2\text{O}} + A_{\text{XY}}[\text{urea}]$ . The units are kilocalories per mole per molar. <sup>c</sup>  $C_{\text{mXY}}$  is the concentration of urea at the midpoint of the transition between conformations X and Y, defined by  $C_{\text{mXY}} = -\Delta G_{\text{XY}}^{\text{H}_2\text{O}}/A_{\text{XY}}$ . The units are molar. <sup>d</sup>  $\Delta\Delta G_{\text{XY}}$  is the difference in the free energy changes between EC  $\alpha$  subunit and the ST  $\alpha$  subunit or the hybrid  $\alpha$  subunits,  $\Delta\Delta G_{\text{XY}} = \Delta G_{\text{XY}}(\text{ST or hybrid } \alpha \text{ subunit}) - \Delta G_{\text{XY}}(\text{EC } \alpha \text{ subunit})$ , calculated at 2.62 M urea for the N  $\leftrightarrow$  I transition and at 4.18 M urea for the I  $\leftrightarrow$  U transition.

where  $Z = (\epsilon_{\text{I}} - \epsilon_{\text{N}})/(\epsilon_{\text{U}} - \epsilon_{\text{N}})$ . A value of 0.59 for  $Z$  which has been used previously for EC  $\alpha$  subunit (Beasty et al., 1986) gave satisfactory fits for all the proteins in this study. The free energies of unfolding were assumed to depend linearly on the urea concentration (Schellman, 1978), resulting in an exponential dependence for the equilibrium constants:

$$K_{\text{XY}} = \exp[-(\Delta G_{\text{XY}}^{\text{H}_2\text{O}} - A[\text{urea}])/RT]$$

where  $\Delta G_{\text{XY}}^{\text{H}_2\text{O}}$  is the free energy difference between species X and Y in the absence of denaturant,  $A$  is a parameter that reflects the cooperativity of the unfolding transition, and  $R$  and  $T$  are the gas constant and the absolute temperature, respectively. These fits were performed with a nonlinear least-squares fitting program, NLIN (SAS Institute, 1985).

The transient responses to changes in urea concentration were fit to one or a sum of exponentials by the nonlinear least-squares fitting program NLIN (SAS Institute, 1985), and the amplitudes and relaxation times for each kinetic phase were obtained (Beasty et al., 1986). Data collection and fitting were simplified by interfacing the spectrophotometer to a Digital PDP 11/23 computer.

## RESULTS

**Nomenclature.** The  $\alpha$  subunits of tryptophan synthase from *Escherichia coli* (EC  $\alpha$  subunit) and *Salmonella typhimurium* (ST  $\alpha$  subunit) differ at 40 positions in the sequence of 268 amino acids (Schneider et al., 1981). Of these differences, 15 occur in the amino folding unit, residues 1-188, and 25 in the carboxyl folding unit, residues 189-268 (Higgins et al., 1979).<sup>2</sup> These replacements do not interfere significantly with the ability of the ST  $\alpha$  subunit to form an active  $\alpha_2\beta_2$  complex with the  $\beta_2$  subunit of tryptophan synthase from *Escherichia coli* (Schneider et al., 1981). Five hybrid  $\alpha$  subunits were previously constructed by Schneider et al. (1981), who used random homologous recombination of the genes for the two

parent proteins. The three-dimensional structure of these hybrid proteins is presumed to be closely similar to the wild-type proteins because they are all capable of activating the  $\beta_2$  subunit of tryptophan synthase from *Escherichia coli*, and they are all recognized by antiserum to the EC  $\alpha$  subunit (Schneider et al., 1981). The regions in which recombination occurred and the amino acid sequences of the wild-type proteins and the hybrid proteins are shown in Figure 1.

The construction of the hybrid proteins was done in such a way that the amino portion is from *Salmonella typhimurium* while the carboxyl portion is from *Escherichia coli*. The nomenclature introduced by Schneider et al. (1981) to describe these hybrid proteins is that the portion from *Salmonella typhimurium* is named first and is designated by the number of amino acid replacements that occur in that region. The portion from *Escherichia coli* is named second and is designated by the number of replacements in that region. Thus, hybrid 15-25 consists of the first 184 residues from the ST  $\alpha$  subunit, where there are 15 differences with the EC  $\alpha$  subunit, and the last 84 residues from the EC  $\alpha$  subunit, where there are 25 differences from the ST  $\alpha$  subunit. The other hybrids in this set are 14-26, 12-28, 8-32, and 6-34.

**$\alpha$  Subunit from *Salmonella typhimurium*.** The equilibrium and kinetic properties of the urea-induced unfolding and refolding reactions at pH 7.8, 25 °C for the ST  $\alpha$  subunit were monitored by difference ultraviolet spectroscopy at 286 nm (Figure 2A,B). For comparative purposes, the results for the EC  $\alpha$  subunit are also shown. An acceptable fit of the equilibrium data (Figure 2A) requires a three-state model involving the native form (N), a stable intermediate (I), and an unfolded form (U) (Matthews & Crisanti, 1981). Previous studies on the EC  $\alpha$ -subunit protein have shown that the intermediate has a folded amino region, residues 1-188, and an unfolded carboxyl region, residues 189-268 (Miles et al., 1982). The N  $\leftrightarrow$  I transition occurs between 1.5 and 3 M urea and the I  $\leftrightarrow$  U transition between 3 and 5.5 M urea. The effect of 1 or more of the 40 amino acid replacements is to slightly stabilize the N  $\leftrightarrow$  I transition and destabilize the I  $\leftrightarrow$  U transition in the ST  $\alpha$  subunit, relative to the transitions in the EC  $\alpha$  subunit.

Quantitative estimates of the stability and the transition midpoints can be obtained by fitting these data to a three-state model as described under Experimental Procedures. Assuming a linear dependence of the free energy of unfolding on the urea concentration (Schellman, 1978), the free energy difference between N and I in the absence of urea is  $6.1 \pm 0.5$  kcal mol<sup>-1</sup>, and that between I and U is  $6.2 \pm 0.7$  kcal mol<sup>-1</sup> for the ST  $\alpha$  subunit (Table I). The concentration of urea at the midpoints of these two transitions is 2.79 and 3.93 M, respectively.

<sup>2</sup> On the basis of previous studies (Miles et al., 1982; Beasty & Matthews, 1985), the 1-188 segment has been referred to as the *amino domain* and 189-268 as the *carboxyl domain*. This designation was deemed appropriate because the regions can be separated by limited proteolysis and demonstrate independent folding by both equilibrium and kinetic criteria. A preliminary report of the X-ray structure of the ST  $\alpha$  subunit (Hyde et al., 1987) shows that the  $\alpha$  subunit has an eight-stranded  $\alpha/\beta$  barrel motif and appears to be a single structural entity. Region 1-188 contains six contiguous  $\beta$  strands and their associated  $\alpha$  helices, and region 189-268 has two contiguous strands and their associated helices. Because the term domain has usually been associated with an independent, identifiable structural region, the segment from 1 to 188 will be referred to in this paper as the *amino folding unit* and that from 189 to 268 as the *carboxyl folding unit*.

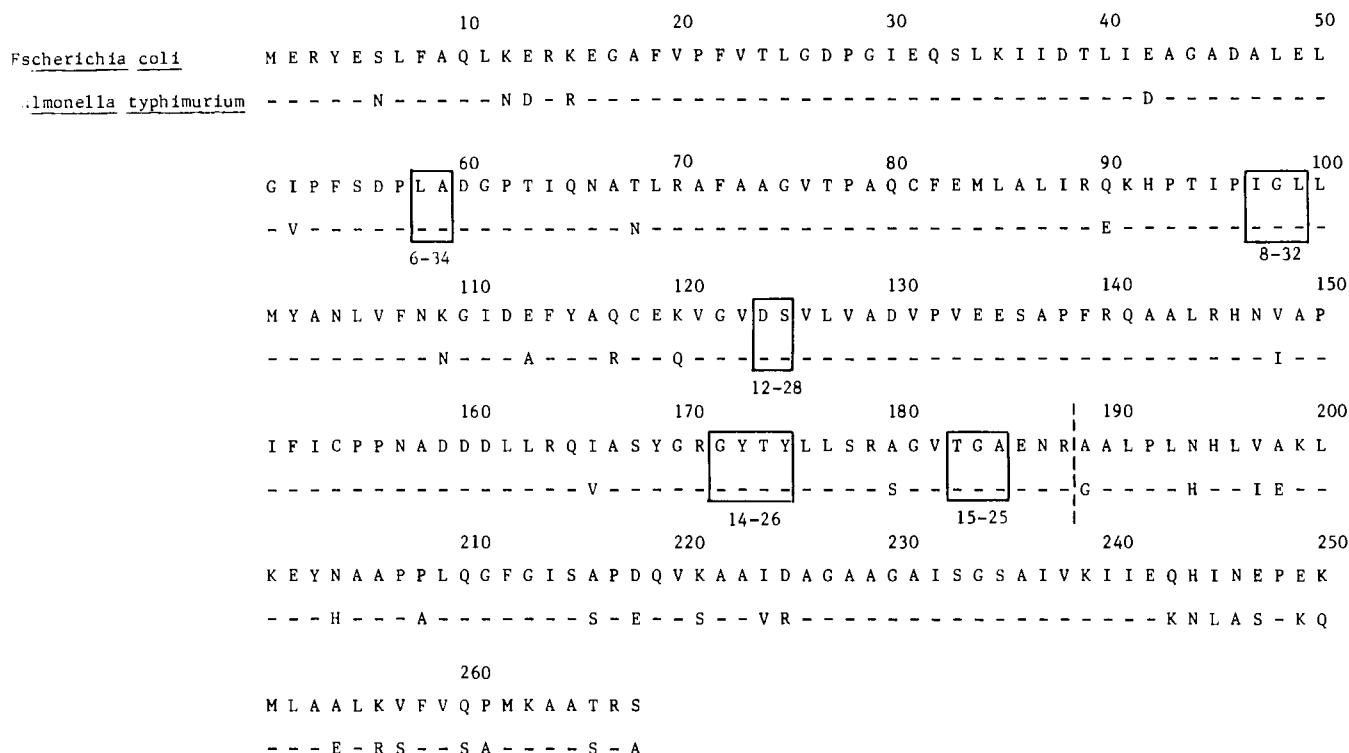


FIGURE 1: Amino acid sequences of EC  $\alpha$  subunit and ST  $\alpha$  subunit (Schneider et al., 1981). Dashes indicate identical amino acids in both sequences. Boxes indicate the regions where recombination has occurred for the respective hybrid  $\alpha$  subunits. In the hybrid  $\alpha$  subunits, the amino-terminal region is always derived from ST  $\alpha$  subunit and the carboxyl-terminal region from EC  $\alpha$  subunit. The vertical dashed line indicates the proposed junction between the amino and carboxyl folding units (Miles et al., 1982).

For comparison, stabilities and transition midpoints for the EC  $\alpha$  subunit are included in Table I. Although there are clear differences between EC and ST  $\alpha$  subunits in the midpoints of the two transitions, the extrapolations required to obtain values for the free energy difference in the absence of urea lead to errors which are comparable to the differences which are due to the amino acid replacements (Table I).

The free energy differences for each transition can also be compared at the urea concentrations corresponding to the midpoints of the  $N \leftrightarrow I$  and  $I \leftrightarrow U$  transitions for the EC  $\alpha$  subunit. This procedure, which has been described in detail elsewhere (Beasty et al., 1986), eliminates the extrapolation required to estimate the stability in the absence of denaturant. The differences in the free energies of the  $N \leftrightarrow I$  and  $I \leftrightarrow U$  transitions between the EC and ST  $\alpha$  subunits calculated by this procedure are also shown in Table I ( $\Delta\Delta G_{NI}$  and  $\Delta\Delta G_{IU}$ ). According to this procedure, ST  $\alpha$  subunit is 0.4 kcal mol<sup>-1</sup> more stable than EC  $\alpha$  subunit in the  $N \leftrightarrow I$  transition and 0.4 kcal mol<sup>-1</sup> less stable than EC  $\alpha$  subunit in the  $I \leftrightarrow U$  transition. Note that these estimates are made in solvent conditions which contain the standard buffer and 2.62 M urea for the  $N \leftrightarrow I$  transition and 4.18 M urea for the  $I \leftrightarrow U$  transition.

Kinetic studies of the folding of the ST  $\alpha$  subunit revealed a single slow phase in unfolding and two slow phases in refolding. The presence of additional, faster phases in refolding is deduced from the observation that the sum of the amplitudes of the two slow phases does not account for the total change in extinction coefficient expected from the equilibrium studies (data not shown). Attempts to directly measure the faster relaxation times by stopped-flow techniques were frustrated by transient aggregation effects.

The urea dependence for the observable relaxation times of the ST  $\alpha$  subunit is qualitatively similar to the dependence for the EC  $\alpha$  subunit (Figure 2B). The unfolding relaxation time decreases logarithmically as the urea concentration in-

Table II: Average Relative Amplitudes for the Slow-Folding Phases in EC, ST, and Hybrid  $\alpha$ -Subunit Proteins

| $\alpha$ subunit | [urea] range (M) | % sucrose<br>(w/v) | $a_1^a$ | $a_2^b$ |
|------------------|------------------|--------------------|---------|---------|
| EC               | 0.6-1.7          | 0                  | 0.26    | 0.74    |
| EC               | 0.6-2.0          | 20                 | 0.40    | 0.60    |
| ST               | 0.6-1.7          | 0                  | 0.30    | 0.70    |
| ST               | 0.7-1.0          | 20                 | 0.28    | 0.72    |
| 8-32             | 0.6-1.0          | 20                 | 0.27    | 0.73    |
| 12-28            | 0.6-1.0          | 20                 | 0.31    | 0.69    |
| 14-26            | 0.6-1.0          | 20                 | 0.19    | 0.81    |
| 15-25            | 0.6-1.0          | 20                 | 0.23    | 0.77    |

<sup>a</sup> Average relative amplitude for the  $\tau_1$  reaction. <sup>b</sup> Average relative amplitude for the  $\tau_2$  reaction.

creases. In refolding, the relaxation time for the slower of the two observable phases is independent of the urea concentration. The behavior of the faster phase is more complex: between 3 and 1.5 M urea, the relaxation time for this phase decreases logarithmically with decreasing urea concentration, while below 1.5 M urea it becomes independent of the urea concentration. This pattern can be understood in terms of a previously proposed kinetic model for the EC  $\alpha$  subunit involving native, intermediate, and unfolded forms (Beasty et al., 1986) (Figure 3).

The similarity in the folding of the ST and EC  $\alpha$  subunits is also observed in the amplitudes of the two observed phases in refolding. Like the EC  $\alpha$  subunit, the amplitudes for the ST  $\alpha$  subunit are independent of the final urea concentration in the native base-line region (data not shown). The average relative amplitudes of the slower and faster of the two phases for ST  $\alpha$  subunit are 0.30 and 0.70, respectively. These values are nearly identical with those for EC  $\alpha$  subunit, 0.26 and 0.74, respectively (Table II). This result and that for the urea dependence of the relaxation times described above show that the ST  $\alpha$  subunit follows the same folding mechanism as the homologue from *E. coli*.

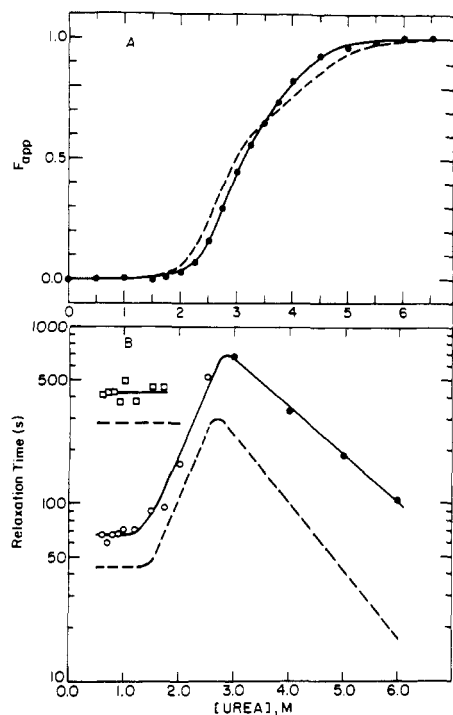


FIGURE 2: (A) Dependence of the apparent fraction unfolded,  $F_{app}$ , on the urea concentration for the ST  $\alpha$  subunit (●) at pH 7.8, 25 °C. The data were derived from difference ultraviolet absorbance studies at 286 nm as described under Experimental Procedures. The line indicates a fit to a three-state model as described in the text. The transition curve for the EC  $\alpha$  subunit (---) under the same conditions is shown for comparison. (B) Semilogarithmic plot of the urea concentration dependence of the relaxation times for the single phase in unfolding (●) and the fast (○) and slow (□) phases in refolding for the ST  $\alpha$  subunit at pH 7.8, 25 °C. The solid line is added to guide the eye and does not represent a fit of the data to an equation. The dependence for the EC  $\alpha$  subunit (---) is shown for comparison.

The quantitative differences observed in the relaxation times for the EC and ST  $\alpha$  subunit demonstrate that 1 or more of the 40 amino acid replacements play a significant role in folding and/or stability. Unfolding is slowed by a factor of 6 at 6 M urea; for refolding at 2 M urea, the  $I_3 \rightarrow N$  ( $\tau_3$ ) reaction (Figure 2) slows by a factor of 2. Below 1.5 M urea, where the relaxation times for both phases are urea independent and the  $I_1 \rightarrow I_2$  and  $I_2 \rightarrow I_3$  reactions (Figure 3) are rate limiting, the average value of the relaxation time for the  $I_1 \rightarrow I_2$  reaction ( $\tau_1$ ) is 420 s and that for the  $I_2 \rightarrow I_3$  reaction ( $\tau_2$ ) is 67 s. The relaxation times for the corresponding phases for the EC  $\alpha$  subunit are 282 and 44 s, respectively (Matthews et al., 1983).

**$\alpha$ -Subunit Hybrids.** The equilibrium unfolding transitions for the five hybrid  $\alpha$  subunits, 6–34, 8–32, 12–28, 14–26, and 15–25, were also monitored by difference ultraviolet spectroscopy at 286 nm (Figure 4). All five hybrid proteins are destabilized to a similar extent in the  $N \leftrightarrow I$  and  $I \leftrightarrow U$  transitions relative to both parent proteins. The transition midpoints and free energies of unfolding in the absence of urea are shown in Table I. Also shown are the relative stabilities of the hybrids at the urea concentrations corresponding to the midpoints of the  $N \leftrightarrow I$  and  $I \leftrightarrow U$  transitions for the EC  $\alpha$  subunit. According to the latter basis of comparison, the hybrids show a decrease in stability ranging from 1.1 to 2.2 kcal mol<sup>-1</sup> for the  $N \leftrightarrow I$  transition and from 0.9 to 1.6 kcal mol<sup>-1</sup> for the  $I \leftrightarrow U$  transition. Considering the estimated errors, the differences in stability between the various hybrids may not be significant.

The results of kinetic studies on unfolding and refolding of the five hybrids are shown in Figure 5A,B. All five hybrids

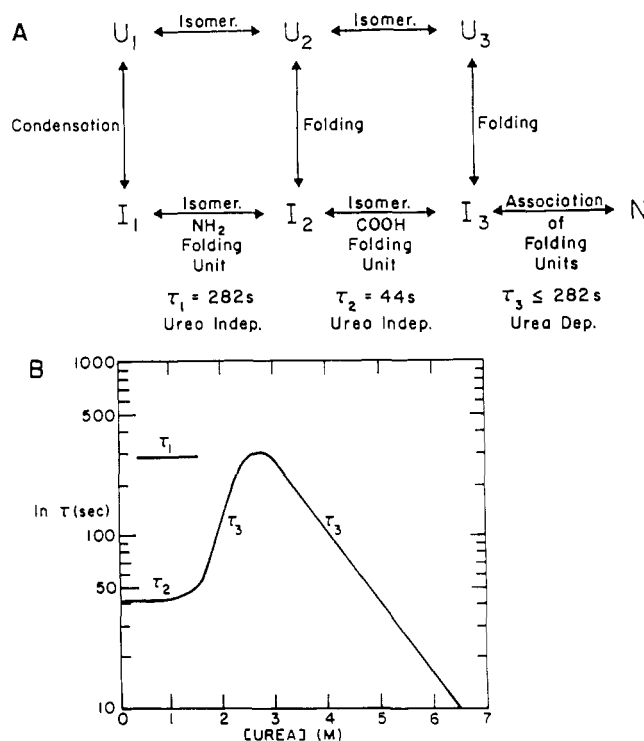


FIGURE 3: (A) Kinetic model for the folding of EC  $\alpha$  subunit at pH 7.8, 25 °C.  $U_1$  and  $U_2$  are two kinetically distinguishable unfolded forms that are separated by a slow-folding reaction in the amino folding unit, residues 1–188 (Crisanti & Matthews, 1981; Hurle & Matthews, 1987).  $U_3$  is a third type of unfolded form that is thought to differ from  $U_2$  by the state of isomerization of an X-Pro peptide bond in the carboxyl folding unit, residues 189–268 (Crisanti & Matthews, 1981; Hurle & Matthews, 1987).  $I_1$ ,  $I_2$ , and  $I_3$  are a series of intermediates that appear during refolding. These intermediates presumably differ for the same reason as the unfolded forms. The conversion of  $I_3$  to  $N$  is limited by association of the two folding units or perhaps by another type of molecule-wide reaction (Beasty et al., 1986). (B) A representative plot of  $\ln \tau$  versus the urea concentration for EC  $\alpha$  subunit, showing the assignment of the observed relaxation times to specific steps in the folding model. The change in the urea dependence of the faster refolding reaction at  $\sim 1.5$  M urea is thought to reflect a change in the process which is rate limiting for the folding of  $I_2$ : Above  $\sim 1.5$  M urea, the  $I_3 \rightarrow N$  step ( $\tau_3$ ) is rate limiting, while below  $\sim 1.5$  M urea  $I_2 \rightarrow I_3$  ( $\tau_2$ ) becomes rate limiting. This change occurs because the  $I_3 \rightarrow N$  reaction accelerates at lower urea concentration while the  $I_2 \rightarrow I_3$  reaction is independent of the urea concentration.

unfold in a single kinetic phase whose relaxation time is within experimental error of that for the EC  $\alpha$  subunit (Figure 5A). For refolding, the relaxation time for the  $I_3 \rightarrow N$  reaction ( $\tau_3$ ) has a similar magnitude for all five hybrids and is longer than that for the EC  $\alpha$  subunit. Not apparent for any of the hybrids is the urea concentration independence of this phase at low urea concentrations (where the  $I_2 \rightarrow I_3$  reaction,  $\tau_2$ , becomes rate limiting) observed in both parent proteins (Figure 2B). The relaxation time for the  $I_1 \rightarrow I_2$  reaction ( $\tau_1$ ) for the 6–34, 8–32, 12–28, and 15–25 hybrids appears to be urea concentration independent and in the same time range as observed for the parent proteins (Figure 2B), although there is a good deal of scatter in some of these measurements (Figure 5B). In contrast, the corresponding relaxation time for the 14–26 hybrid is significantly longer and has become urea concentration dependent; the relaxation time decreases as the final urea concentration decreases (Figure 5B).

To determine if these changes in the urea concentration dependence of the refolding relaxation times for the hybrid proteins mean that the folding mechanism has changed, the folding studies were performed in the presence of 20% sucrose.

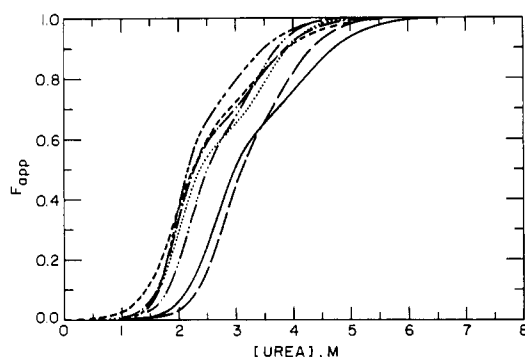


FIGURE 4: Dependence of the apparent fraction unfolded,  $F_{app}$ , on the urea concentration for the EC  $\alpha$  subunit (—), the ST  $\alpha$  subunit (---), and the 6-34 (— · —), 8-32 (·····), 12-28 (--- · ---), 14-26 (·····), and 15-25 (---) hybrid  $\alpha$  subunits at pH 7.8, 25 °C. The data were derived from difference ultraviolet absorbance studies at 286 nm as described under Experimental Procedures. The lines indicate results of the nonlinear least-squares fit of the data to a three-state model as described in the text. To simplify the figure, the data points have not been included; however, the computer fits to the actual data are comparable to that shown in Figure 2A for the ST  $\alpha$  subunit.

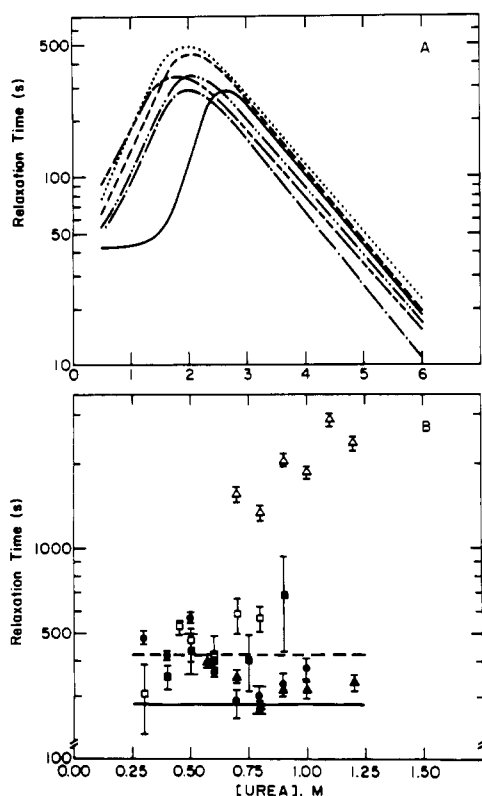


FIGURE 5: (A) Semilogarithmic plot of the urea concentration dependence of the  $\tau_3$  relaxation time for unfolding and refolding of the EC  $\alpha$  subunit (—) and the 6-34 (— · —), 8-32 (·····), 12-28 (--- · ---), 14-26 (·····), and 15-25 (---) hybrid  $\alpha$  subunits at pH 7.8, 25 °C. The data points have not been included in order to simplify the figure. The scatter of the data about the lines is comparable to that shown for the ST  $\alpha$  subunit in Figure 2B. (B) Semilogarithmic plot of the urea concentration dependence of the slower,  $\tau_1$ , relaxation time for refolding of the 6-34 ( $\blacktriangle$ ), 8-32 ( $\bullet$ ), 12-28 ( $\square$ ), 14-26 ( $\triangle$ ), and 15-25 ( $\square$ ) hybrid  $\alpha$  subunits at pH 7.8, 25 °C. The average values for the  $\tau_1$  phase in the EC  $\alpha$  subunit (—) and the ST  $\alpha$  subunit (---) are shown for reference.

Consideration of the previously proposed kinetic model for the folding of the EC  $\alpha$  subunit (Matthews et al., 1983) suggests that these changes could be explained for the  $I_1 \rightarrow I_2$  reaction by the destabilization of intermediate  $I_1$  with respect to unfolded form  $U_1$  and for the  $I_2 \rightarrow I_3$  reaction by the destabilization of the native form and the transition state linking N

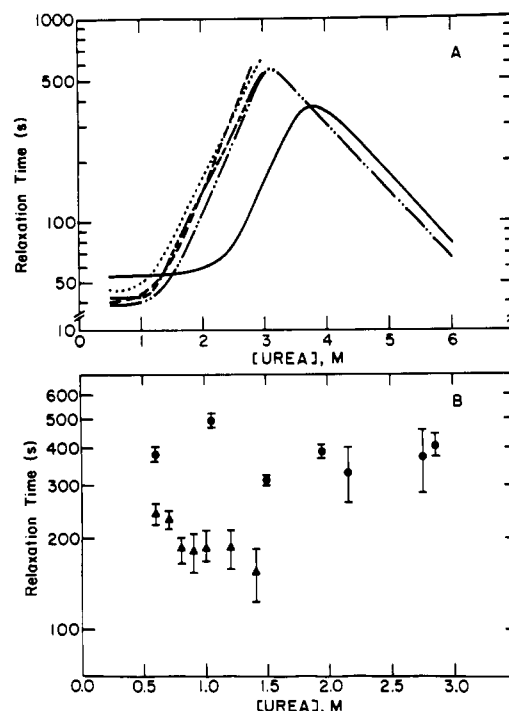


FIGURE 6: (A) Semilogarithmic plot of the urea concentration dependence of the  $\tau_2$  and  $\tau_3$  relaxation times for refolding and the  $\tau_3$  phase for unfolding in the presence of 20% (w/v) sucrose for the EC  $\alpha$  subunit (—) and the 8-32 (— · —) hybrid  $\alpha$  subunit at pH 7.8, 25 °C. Refolding data only are shown for the 12-28 (--- · ---), 14-26 (·····), and 15-25 (---) hybrid  $\alpha$  subunits. (B) Semilogarithmic plot of the urea concentration dependence of the  $\tau_1$  relaxation time for refolding in the presence of 20% (w/v) sucrose for the EC  $\alpha$  subunit ( $\bullet$ ) and the 14-26 ( $\blacktriangle$ ) hybrid  $\alpha$  subunits at pH 7.8, 25 °C.

and  $I_3$  with respect to intermediate  $I_3$ . This possibility can be tested by performing the refolding studies in the presence of sucrose because sucrose stabilizes the folded conformations of proteins relative to more unfolded forms (Lee & Timasheff, 1981).

The effect of 20% sucrose on the  $I_2 \rightarrow I_3$  ( $\tau_2$ ) and  $I_3 \rightarrow N$  ( $\tau_3$ ) reactions in refolding for four of the five hybrids is shown in Figure 6A. The urea-dependent  $\tau_3$  phase for the 8-32, 12-28, 14-26, and 15-25 hybrids shifts to higher urea concentration. This behavior has been observed previously with EC  $\alpha$  subunit (Hurle et al., 1987) and is consistent with an increase in stability of the native form with respect to the intermediate form.

Also apparent for four of the five hybrids is the urea-independent relaxation time ( $\tau_2$ ) below ~1 M urea. The  $\tau_2$  relaxation times for the four hybrids in 20% sucrose range from 39 to 46 s and are equivalent, within experiment error. For comparison, the effects of 20% sucrose on the refolding kinetics of ST  $\alpha$  subunit were also measured. The  $\tau_2$  reaction for the ST  $\alpha$  subunit approaches a urea-independent value of 47 s at low urea concentrations (data not shown). It is interesting to note that the relaxation times for the  $\tau_2$  reactions for the EC and ST  $\alpha$  subunits become nearly equivalent in sucrose and are not distinguishable from the values for the hybrids. This contrasts with the result in the absence of sucrose where ST  $\alpha$  subunit clearly folds more slowly than EC  $\alpha$  subunit by the  $\tau_2$  pathway (Figure 2B).

The relaxation time for the  $\tau_1$  phase for the 14-26 hybrid decreases significantly in 20% sucrose and becomes comparable to that for the EC  $\alpha$  subunit (Figure 6B). The relaxation time does not show a strong dependence on the urea concentration and, considering the estimated errors, appears to be independent of the urea concentration. The scatter in the  $\tau_1$  data

reflects the difficulty in accurate measurement of slow kinetic phases of low amplitude. The  $\tau_1$  phase usually represents less than 10% of the observed change in extinction coefficient. Therefore, in the presence of sucrose, the refolding kinetics of the hybrids resemble those of the EC and the ST  $\alpha$  subunits.

Further confirmation of the similarities of the folding mechanisms of the parent and hybrid  $\alpha$  subunits can be obtained by comparison of the amplitudes of the two slow-folding reactions for jumps into the native base-line region. The relative amplitudes for the hybrids in the presence of 20% sucrose are independent of the final urea concentration (data not shown) and are rather similar to the values observed for the EC and ST  $\alpha$  subunits (Table II). Because these amplitudes reflect the relative populations of the unfolded forms,  $U_1$  and  $U_2$ , prior to refolding, it is clear that the hybrids do not significantly alter this distribution. The similarities of these amplitudes in the presence and absence of 20% sucrose for each of the parent proteins suggest that the presence of sucrose also does not alter this distribution to any great extent.

## DISCUSSION

The equilibrium and kinetic results for the folding of the ST  $\alpha$  subunit and five *Escherichia coli*-*Salmonella typhimurium* hybrids clearly show that the folding mechanism previously proposed for the EC  $\alpha$  subunit also applies to these proteins. This conclusion supports the hypothesis advanced by Hollecker and Creighton (1983) and by Krebs et al., (1983) that the conservation of the folding mechanism is an important evolutionary pressure. The strength of this conclusion is tempered by the relatively high (85%) degree of conservation between the EC and ST  $\alpha$  subunits. Studies on the rate-limiting steps in the folding of the  $\alpha/\beta$  barrel structures with less or no homology to  $\alpha$  subunit are required to demonstrate whether structural families fold by common mechanisms.

The data on the folding kinetics also provide information on the question of whether the nonconserved residues play a role in one of the rate-limiting steps in folding. According to the previously proposed kinetic model for folding (Matthews et al., 1983), the  $\tau_3$  phase (Figures 2B) reflects the interconversion between the native form, N, and a stable intermediate,  $I_3$  (Figure 3). Unfolding dominates above 3 M urea for EC  $\alpha$  subunit and refolding below 3 M urea. The maximum in the plot of  $\ln \tau$  versus urea concentration closely corresponds to the midpoint in the  $N \leftrightarrow I_3$  equilibrium transition, where unfolding and refolding rate constants are equal.

Inspection of the relaxation data for the five hybrids (Figure 5A) shows that, when compared to the results for EC  $\alpha$  subunit, the unfolding relaxation time ( $N \rightarrow I_3$ ) is not altered. In contrast, the refolding relaxation time ( $I_3 \rightarrow N$ ) is increased by the same amount for all five hybrids. Because the effect of the hybrids is to selectively alter 1 of the relaxation times, it appears that the 15 replacements in the amino folding unit do not affect the rate-limiting step in the  $N \leftrightarrow I_3$  interconversion. Replacements that do affect the rate-limiting step selectively alter, by definition, the energy of the transition state linking N and  $I_3$  and, therefore, both unfolding and refolding relaxation times (Beasty et al., 1986). This conclusion must be regarded as tentative because more than one amino acid is varied when the hybrids are compared with the EC  $\alpha$  subunit. Single amino acid replacements are required to ascertain the validity of this conclusion.

Comparison of the relaxation data for the EC and ST  $\alpha$  subunits (Figure 2B) shows that both unfolding and refolding relaxation times for the  $N \leftrightarrow I_3$  reaction are altered. As described above, this result implies an effect on the rate-limiting step linking N and  $I_3$ . Considering the results for the

hybrids, the basis for this effect probably resides in the carboxyl folding unit, residues 189–268.

The similar stability (Figure 4 and Table I) and kinetics of folding (Figure 5A) for the five hybrids shows that the replacements that occur across the 6–34 to 15–25 set, Asn-68  $\rightarrow$  Thr (ST amino acid  $\rightarrow$  EC amino acid), Glu-90  $\rightarrow$  Gln, Asn-109  $\rightarrow$  Lys, Ala-113  $\rightarrow$  Glu, Arg-117  $\rightarrow$  Gln, Gln-120  $\rightarrow$  Lys, Ile-148  $\rightarrow$  Val, Val-166  $\rightarrow$  Ile, and Ser-180  $\rightarrow$  Ala, conserve the relative stabilities of the native, intermediate, and unfolded forms. Because 14–26 and 15–25 differ by a single amino acid (14–26 has Ala at position 180 while 15–25 has Ser), it is clear that this replacement does not alter the stability of the principal intermediates,  $I_2$  and  $I_3$ , with respect to the native and unfolded forms. Pairwise comparisons of the other hybrids cannot lead to similar definitive conclusions because more than one amino acid is changed.

Interestingly, the Ser-180  $\rightarrow$  Ala replacement does decrease the stability of the minor intermediate  $I_1$  with respect to  $U_1$ . This effect is apparent in the  $\tau_1$  phase in refolding for the 14–26 hybrid which becomes significantly slower and dependent on the final urea concentration (Figure 5B). A similar effect has been seen with two single amino acid mutations in the EC  $\alpha$  subunit near position 180, Tyr-175  $\rightarrow$  Cys and Thr-183  $\rightarrow$  Ile (Hurle et al., 1986; N. B. Tweedy, unpublished results).

Comparison of the equilibrium properties of hybrid 6–34 and the EC  $\alpha$  subunit (Figure 4 and Table I) shows that one or more of the six variable amino acids in the amino terminus, Asn-6  $\rightarrow$  Ser, Asn-12  $\rightarrow$  Lys, Asp-13  $\rightarrow$  Glu, Arg-15  $\rightarrow$  Lys, Asp-42  $\rightarrow$  Glu, and Val-52  $\rightarrow$  Ile, contributes significantly to the stability of the protein. Because the stabilities of the  $N \leftrightarrow I$  and  $I \leftrightarrow U$  transitions for the 6–34 hybrid are not a weighted average of those for the EC and ST  $\alpha$  subunits ( $\Delta\Delta G_{NI}$  and  $\Delta\Delta G_{IU}$  in Table I), one or more of the six replacements in this region must have a nonadditive effect on stability (Hurle et al., 1986).

This conclusion contradicts that reached by Yutani et al. (1984) in comparing the stabilities of the EC, ST, and 14–26 hybrid  $\alpha$  subunits. This contradiction appears to arise from a different method of estimating the stabilities using chemical denaturants and not from a fundamental difference in the folding mechanism. Yutani et al. (1984) estimated the stability to guanidine hydrochloride induced unfolding by extrapolation to zero molar denaturant, a procedure that can lead to uncertainty both from the method of extrapolation and from statistical errors in fitting.

Cupo and Pace (1983) have suggested that the most accurate method for comparing the stabilities of a series of chemically modified proteins is to compare the stabilities at the midpoint of the transition for the unmodified protein, as was done in the present paper. Applying this method to the data in Figure 1 of Yutani et al. (1984), it is apparent that the  $N \leftrightarrow I$  reaction for the 14–26 hybrid is less stable to guanidine hydrochloride induced unfolding than either parent protein. This conclusion is in agreement with the present result using urea and is consistent with a common unfolding mechanism for the two chemical denaturants.

The apparent additivity reported by Yutani et al. (1984) for the thermal transitions of EC, ST, and the 14–26 hybrid  $\alpha$  subunits cannot be compared to the results for urea or guanidine hydrochloride because the thermal transition follows a two-state reaction (Matthews et al., 1980) and does not involve a stable intermediate.

For the  $N \leftrightarrow I$  transition, the kinetic manifestation of the increase in stability for the EC  $\alpha$  subunit relative to the 6–34

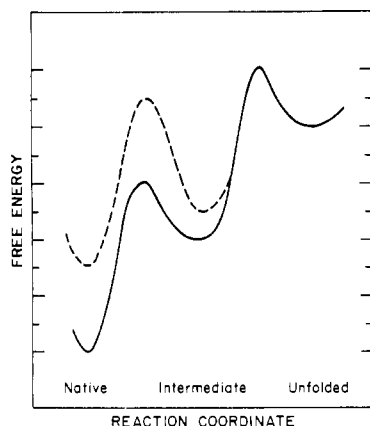


FIGURE 7: Hypothetical reaction coordinate diagrams for the EC  $\alpha$  subunit (—) and the 6-34 hybrid  $\alpha$  subunit (---). The diagrams have been arbitrarily aligned by equating the free energies of the unfolded forms and exaggerated to emphasize the selective effects of the hybrid on the  $I \rightarrow N$  refolding reaction; the rate of the  $N \rightarrow I$  unfolding reaction is the same for the 6-34 hybrid and EC  $\alpha$  subunits. The energies of the transition states for the  $I \leftrightarrow U$  reaction have been arbitrarily equated. No data are available on the rates of the  $I \rightarrow U$  or  $U \rightarrow I$  reactions.

hybrid is a decreased relaxation time for refolding (Figure 5A). These results can be summarized in terms of a reaction coordinate diagram (Figure 7). For the 6-34 hybrid  $\alpha$  subunit, the observed reduced stabilities for the  $N \leftrightarrow I$  and  $I \leftrightarrow U$  transitions are reflected in decreased free energy differences both between  $N$  and  $I$  and between  $I$  and  $U$ . The selective effect on the refolding relaxation time implies that the effect of the six amino acid replacements on the free energy of the transition state between  $N$  and  $I$  must be very similar to that on the free energy of the native conformation.

Comparison of the stabilities of the EC, ST, and 15-25 hybrid  $\alpha$  subunits (Figure 4 and Table I) suggests that the free energy of unfolding of the carboxyl folding unit (the  $N \leftrightarrow I$  equilibrium transition) depends upon whether that sequence is fused to an EC or an ST amino sequence ( $\Delta\Delta G_{NI}$  in Table I). Compared to the EC  $\alpha$  subunit,  $\Delta\Delta G_{NI}$  decreases by  $1.4 \pm 0.9$  kcal mol<sup>-1</sup> for the 15-25 hybrid protein, and the midpoint of the  $N \leftrightarrow I$  transition decreases from  $2.62 \pm 0.02$  M urea to  $1.92 \pm 0.18$  M urea (Table I). These results may reflect either a different conformation for the carboxyl folding unit in the hybrid or a decreased affinity between the amino and carboxyl folding units in the hybrid.

For the  $I \leftrightarrow U$  transition, which corresponds to the unfolding of the amino folding unit, the errors in  $\Delta\Delta G_{IU}$  for the ST  $\alpha$  subunit and the 15-25 hybrid  $\alpha$  subunit (Table I) preclude a definitive statement concerning the relative stability of the ST amino folding unit fused to either an ST carboxyl folding unit or an EC carboxyl folding unit. Previous studies have shown that the replacement of Gly-211, in the carboxyl folding unit, with Glu alters the free energy change for the  $I \leftrightarrow U$  transition (Beasty et al., 1986). Because this effect disappears in the presence of 0.5 M NaCl, it is thought to be due to electrostatic interactions between the folded amino folding unit and the unfolded carboxyl folding unit.

The urea-independent relaxation times observed for refolding under strongly native conditions for all seven  $\alpha$ -subunit proteins show that these rate-limiting reactions are also conserved. A previous study (Crisanti & Matthews, 1981) suggested that the  $\tau_1$  and  $\tau_2$  reactions might well be cis-trans isomerizations at X-Pro peptide bonds (Brandts et al., 1975); however, recent results now suggest that the  $\tau_1$  reaction is another type of isomerization process (Hurle & Matthews, 1987).

The relaxation time for Pro isomerization reactions is known

to depend on the local amino acid sequence, particularly on the preceding amino acid (Brandts et al., 1975; Grathwohl & Wuthrich, 1981). Inspection of the sequence in the carboxyl folding unit, with which the  $\tau_2$  phase has been previously associated (Crisanti & Matthews, 1981), shows that the EC  $\alpha$  subunit has, in addition to the four Pro residues found in the ST  $\alpha$  subunit, Pro residues at positions 208 and 261. The observation that the  $I_2 \rightarrow I_3$  reaction ( $\tau_2$ ) is more rapid in EC than ST  $\alpha$  subunit suggests that neither of the two additional Pro residues in EC  $\alpha$  subunit is limiting the folding. Two of the remaining four Pro residues show amino acid replacements at the preceding positions: Pro-217 where Ser-216 (EC)  $\rightarrow$  Ala (ST) and Pro-248 where Ser-247 (EC)  $\rightarrow$  Glu (ST). Mutagenesis is required to determine if one of these Pro residues is responsible for this rate-limiting step in folding.

The conclusion that only 1 of the 19 X-Pro peptide bonds in the  $\alpha$  subunit is capable of retarding the rate of folding is contradictory to the original hypothesis of Brandts et al. (1975). However, it is in agreement with observations in other proteins (Cook et al., 1979; Rudolph et al., 1983) and has led to the concept of "essential" and "nonessential" X-Pro peptides bonds (Jullien & Baldwin, 1981; Levitt, 1980). The former class is capable of slowing the folding reaction while the latter is not. Nall (1985) has discussed the proline isomerization problem in a recent review.

The results presented in this paper on the folding of a series of homologous proteins support the hypothesis that the conservation of the folding mechanism is an important evolutionary pressure. Amino acid replacements that dramatically slow the folding process might be selected against because the unfolded form or partially folded intermediates would be susceptible to proteolytic degradation (Goldberg & Dice, 1974). Also, the lower solubility of unfolded or partially folded forms in an intracellular environment could lead to competing aggregation reactions that would reduce the yield of the native, functional form (Marston, 1986). Single amino acid replacements are required for a rigorous test of the validity of this hypothesis. The amino acid positions at which EC and ST  $\alpha$  subunits differ provide a starting point for such experiments.

#### ACKNOWLEDGMENTS

We gratefully acknowledge the assistance of Mark Hurle in data analysis and for a critical review of the manuscript, Neil Tweedy for a critical review of the manuscript, and Gail Feldman for typing the manuscript.

#### REFERENCES

- Ahmed, S. A., Miles, E. W., & Davies, D. R. (1985) *J. Biol. Chem.* 260, 3716.
- Bajaj, M., & Blundell, T. (1984) *Annu. Rev. Biophys. Bioeng.* 13, 453.
- Beasty, A. M., Hurle, M. R., Manz, J. T., Stackhouse, T. S., Onuffer, J. J., & Matthews, C. R. (1986) *Biochemistry* 25, 2965.
- Brandts, J. F., Halvorsen, H. R., & Brennan, M. (1975) *Biochemistry* 14, 4953.
- Cook, K. H., Schmid, F. X., & Baldwin, R. L. (1979) *Proc. Natl. Acad. Sci. U.S.A.* 76, 6157.
- Crisanti, M. M., & Matthews, C. R. (1981) *Biochemistry* 20, 2700.
- Cupo, J. F., & Pace, C. N. (1983) *Biochemistry* 22, 2654.
- Goldberg, A. L., & Dice, J. F. (1974) *Annu. Rev. Biochem.* 43, 835.
- Grathwohl, C., & Wuthrich, K. (1981) *Biopolymers* 20, 2623.
- Herskovits, T. T. (1967) *Methods Enzymol.* 11, 748.

- Higgins, W., Fairwell, T., & Miles, E. W. (1979) *Biochemistry* 18, 4827.
- Hollecker, M., & Creighton, T. E. (1983) *J. Mol. Biol.* 168, 409.
- Hurle, M. R., & Matthews, C. R. (1987) *Biochim. Biophys. Acta* 913, 179.
- Hurle, M. R., Tweedy, N. B., & Matthews, C. R. (1986) *Biochemistry* 25, 6356.
- Hurle, M. R., Michelotti, G. A., Crisanti, M. M., & Matthews, C. R. (1987) *Proteins (3rd Ed.)* 2, 54.
- Hyde, C. C., Padlan, E. A., Ahmed, S. A., Miles, E. W., & Davies, D. R. (1987) *Fed. Proc., Fed. Am. Soc. Exp. Biol.* 46, 2215A.
- Jullien, M., & Baldwin, R. L. (1981) *J. Mol. Biol.* 145, 265.
- Kirschner, K., Wiskocil, R. L., Loehn, M., & Rezeau, L. (1975) *Eur. J. Biochem.* 60, 513.
- Krebs, H., Schmid, F. X., & Jaenicke, R. (1983) *J. Mol. Biol.* 169, 619.
- Lang, K., Wrba, A., Krebs, H., Schmid, F. X., & Beintema, J. J. (1986) *FEBS Lett.* 204, 135.
- Lee, J. C., & Timasheff, S. N. (1981) *J. Biol. Chem.* 256, 7193.
- Levitt, M. (1980) *J. Mol. Biol.* 145, 251.
- Marston, F. A. O. (1986) *Biochem. J.* 240, 1.
- Matthews, C. R., & Crisanti, M. M. (1981) *Biochemistry* 20, 784.
- Matthews, C. R., Crisanti, M. M., Manz, J. T., & Gepner, G. L. (1983) *Biochemistry* 22, 1445.
- Miles, E. W., Yutani, K., & Ogasahara, K. (1982) *Biochemistry* 21, 2586.
- Miles, E. W., Bauerle, R., & Ahmed, S. A. (1987) *Methods Enzymol.* (in press).
- Nall, B. T. (1983) *Comments Mol. Cell. Biophys.* 3, 123.
- Rossmann, M., & Argos, P. (1981) *Annu. Rev. Biochem.* 50, 497.
- Rudolph, R., Bohrer, M., & Fischer, S. (1983) *Eur. J. Biochem.* 131, 383.
- SAS Institute Inc. (1985) *SAS User's Guide: Statistics, Version 5 Edition*, p 575, SAS Institute Inc., Cary, NC.
- Schellman, J. A. (1978) *Biopolymers* 17, 1305.
- Schmid, F. X., Grafl, R., Wrba, A., & Beintema, J. J. (1986) *Proc. Natl. Acad. Sci. U.S.A.* 83, 872.
- Schneider, W. P., Nichols, B. P., & Yanofsky, C. (1981) *Proc. Natl. Acad. Sci. U.S.A.* 78, 2169.
- Yutani, K., Ogasahara, K., Suzuki, M., & Sugino, Y. (1979) *J. Biochem. (Tokyo)* 95, 915.
- Yutani, K., Sato, T., Ogasahara, K., & Miles, E. W. (1984) *Arch. Biochem. Biophys.* 229, 448.