# Structure and Stability of an Early Folding Intermediate of *Escherichia coli trp* Aporepressor Measured by Far-UV Stopped-Flow Circular Dichroism and 8-Anilino-1-naphthalene Sulfonate Binding†

Craig J. Mann and C. Robert Matthews*

*Department of Chemistry and the Center for Biomolecular Structure and Function, The Pennsylvania State University, University Park, Pennsylvania 16802*

*Received October 26, 1992; Revised Manuscript Received March 9, 1993*

ABSTRACT: The refolding kinetics of *Escherichia coli trp* aporepressor were monitored using stopped-flow far-ultraviolet circular dichroism and 8-anilino-1-naphthalene sulfonate fluorescence spectroscopy. Significant gains in secondary structure and the development of hydrophobic surface, respectively, were observed within the dead time of mixing (4–5 ms). These initial increases, or burst phase amplitudes, plotted as a function of final urea concentration, exhibited sigmoidal, coincident unfolding transition curves. The transition curves were fit to a two-state model, and the resulting free energies of folding in the absence of denaturant were found to be similar (~3.3 kcal/mol). Three subsequent slow refolding phases exhibited relaxation times and amplitudes similar to those previously observed for tryptophan fluorescence [Gittelman, M. S., & Matthews, C. R. (1990) *Biochemistry 29*, 7011–7021]. These results support the proposals that a stable, monomeric intermediate is rapidly formed during the folding of *trp* aporepressor and that this species contains a significant amount of secondary structure and hydrophobic surface. This early intermediate is then processed through folding and association reactions that result in the formation of the remaining secondary, tertiary, and quaternary structure.

The quaternary structure found in oligomeric proteins adds another layer of complexity to the problem of determining the mechanism by which the amino acid sequence directs the rapid and efficient folding to the native conformation. The development of this higher order structure raises several issues that are not encountered in the folding of monomeric proteins: (1) Do the individual subunits acquire structure prior to association? (2) If the subunits do fold, in what time range does this happen and is the structure similar to that in the native conformation? (3) At what stage does the association reaction(s) between subunits occur? and (4) Does folding continue after the association event?

These questions have been addressed in the past by a number of studies on oligomeric protein folding reactions [for reviews, see Jaenicke (1987, 1991)]. For example, the folding of dimeric cytoplasmic malate dehydrogenase proceeds through two parallel folding pathways that include monomeric and dimeric intermediates, respectively, ultimately leading to active native dimers (Rudolph et al., 1986). Garel and co-workers (Le Bras et al., 1989) have proposed a folding mechanism for tetrameric phosphofructokinase that contained monomeric as well as stable dimeric intermediates. Both of these studies focused on slow folding and association reactions and only briefly addressed questions that involve the structure and/or stability of early folding intermediates.

Goldberg and co-workers (Zetina & Goldberg, 1980, 1982; Blond & Goldberg, 1985) have examined the folding of the dimeric $\beta_2$ subunit of tryptophan synthase from a variety of perspectives. These authors have proposed a sequential folding scheme that contained two monomeric intermediates followed by two dimeric intermediates before ultimately leading to the dimeric native species. As in the cases of malate dehydrogenase and phosphofructokinase, folding occurs before and after the association event. More recent experiments have probed the early events in the folding of the $\beta_2$ subunit using stopped-flow circular dichroism (SF-CD)[1] and 8-anilino-1-naphthalene sulfonate (ANS) spectroscopy (Goldberg et al., 1990). A monomeric intermediate appears in the 0–5-s time range and has the characteristics of a molten globule species (Goldberg et al., 1990). This intermediate has at least half of the native secondary structure and is capable of binding the hydrophobic ANS dye in nonpolar regions which are eventually buried in the native conformation. Binding studies with a monoclonal antibody to the native conformation demonstrated that a native antigenic determinant appeared after the formation of the molten globule species, but well before the dimerization reaction leading to the native form. The challenge with the $\beta_2$ subunit is that the overall complexity of the kinetic response during folding has made the development of an actual folding mechanism very difficult.

In contrast, *Escherichia coli trp* aporepressor (TR) appears to be a relatively simple system in which to study the formation of secondary, tertiary, and quaternary structure in an oligomeric protein. TR is a small, all $\alpha$-helical, dimeric protein composed of two identical monomers of 107 amino acids (Gunsalus & Yanofsky, 1980). Each monomer contains six $\alpha$-helices (Zhang et al., 1987), of which three from each monomer (A, B, C, A′, B′, and C′) are intertwined at the dimer interface to form the interior core of the protein (Figure 1). Helices D and E from each monomer make up the "helix–turn–helix" DNA binding motif common to many prokaryotic and phage DNA-binding proteins [for reviews, see Steitz (1990) and Freemont et al. (1991)]. The last helix, F, packs

† This work was supported by the National Institutes of Health, Postdoctoral Fellowship Award GM13571 (to C.J.M.), by the National Institute of General Medical Sciences, Grant GM23303 (to C.R.M.), and by the National Institutes of Health, Shared Instrumentation Award 1S10RR04953.

* Author to whom correspondence should be addressed.

[1] Abbreviations: ANS, 8-anilino-1-naphthalene sulfonate; CD, circular dichroism; IAA, indole acrylic acid; $Na_2EDTA$, ethlylenediaminetetraacetic acid, disodium salt; $NaDodSO_4$, sodium dodecyl sulfate; SF-CD, stopped-flow circular dichroism; SF-Trp Fl, stopped-flow intrinsic tryptophan fluorescence; TR, *trp* aporepressor.

0006-2960/93/0432-5282$04.00/0 

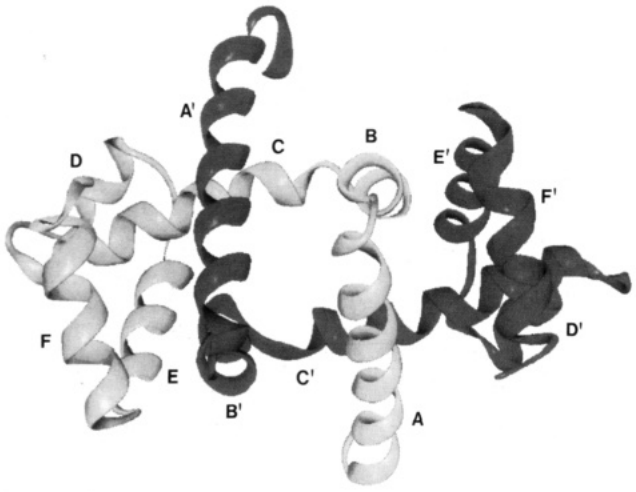



FIGURE 1: Ribbon diagram of TR (Zhang et al., 1987) showing the two intertwined monomers. The six $\alpha$-helices of each subunit are labeled for clarity.

against the core to form a part of the subunit interface region.

Previous work (Gittelman & Matthews, 1990) has shown that the refolding of wild-type TR involves at least three kinetics phases that depend in complex ways on both the protein concentration and the final urea concentration. Between 5.0 and 3.5 M urea, all three refolding phases are urea and protein concentration dependent. This result has been interpreted to mean that the association reactions which form either the native dimer, D, or a pair of dimeric folding intermediates, $D^*_2$ and $D^*_3$, are rate-limiting in the unfolding transition region. At and below 3.0 M urea, these folding reactions accelerate to the point that alternative reactions now become rate-limiting. Empirically, the relaxation times are independent of the protein and the urea concentration. The slowest of the three phases may involve proline isomerization (C. J. Mann, T. R. Garber, and C. R. Matthews, unpublished results). Although the intertwined nature of native conformation would suggest that folding of individual monomers must occur before and after the association reaction, no definitive evidence for folding in a monomer was provided in that study.

In the present study, far-ultraviolet SF-CD was used to probe the development of secondary structure in the millisecond time range during the folding of TR. The binding of the fluorescent dye ANS was also used to follow the formation of hydrophobic surfaces, as well as the development of the corepressor binding site, during the folding of TR. From these and other data, it is proposed that a stable, monomeric intermediate is rapidly formed within the dead time of mixing of the stopped-flow instruments (4–5 ms). This species contains a significant amount of secondary structure and has solvent-exposed hydrophobic surfaces which are ultimately buried in the native conformation. The formation of this early intermediate is followed by folding and association reactions that result in the formation of the remaining secondary, tertiary, and quaternary structure in a concerted process.

## MATERIALS AND METHODS

*Chemicals.* Ultrapure urea was purchased from Schwarz/Mann and used without further purification. ANS was purchased from Molecular Probes, Inc. (Eugene, OR). All other chemicals were reagent grade. The buffer used throughout these experiments contained 10 mM sodium phosphate, pH 7.6, and 0.1 mM $Na_2EDTA$. The temperature was maintained at 25 °C throughout these experiments. The final protein concentration was 0.13 mg/mL (5.3 $\mu$M) as determined by using a molar extinction coefficient of $2.97 \times 10^4$ $M^{-1}$ $cm^{-1}$ at 280 nm (Joachimiak et al., 1983). All aporepressor concentrations are reported in terms of the dimer.

*Protein Purification.* Wild-type *trp* aporepressor was purified as previously described (Chou et al., 1989; Paluh & Yanofsky, 1986). Protein purity was verified by the presence of a single band on Coomassie blue stained $NaDodSO_4$ polyacrylamide gels (Schagger & von Jagow, 1987).

*Intrinsic Tryptophan Fluorescence.* Stopped-flow fluorescence experiments were conducted using a Bio-Logic SFM-3 stopped-flow module equipped with a Bio-Logic modular light source and monochromator. Samples were excited at 295 nm, and emission above 320 nm was detected using a cutoff filter supplied by Bio-Logic. The cuvette path length was 2.0 mm, and the dead time of mixing was 4 ms.

*Circular Dichroism.* Equilibrium CD spectra were determined on an AVIV Associates model 62DS Circular Dichroism Spectrometer. The path length of the cuvette was 2 mm. Stopped-flow measurements were obtained using a Bio-Logic SFM-3 stopped-flow system attached to the AVIV 62DS spectrometer. The path length of the cuvette was 1.5 mm, and the dead time of these experiments was 5 ms. Data collections were made over two time ranges: (1) the first 5 s using an instrumental time constant of 1 ms and (2) the first 200 s using a time constant of 100 ms. To improve the signal-to-noise ratio, 50 shots were averaged for the short time range, and 10 shots were averaged for the long time range. All CD data are reported as mean residue ellipticity by taking 115 g/mol as the mean residue molecular weight of TR.

*Dead-Time Determination.* The dead-times for both the stopped-flow fluorescence instrument and the stopped-flow CD instrument were determined in the absorbance mode using the reduction of 2,6-dichlorophenolindophenol by L-ascorbic acid measured at 524 nm (Tonomura et al., 1978). The dead-time of the fluorescence instrument was 4 ms and that for the CD instrument was 5 ms.

*ANS Binding.* ANS concentration was determined by using a molar extinction coefficient of $6.8 \times 10^3$ $M^{-1}$ $cm^{-1}$ in methanol at 370 nm (Haugland, 1991). Stopped-flow ANS fluorescence experiments were performed in a fashion similar to those described above for intrinsic tryptophan fluorescence with the following exceptions: (1) the final ANS concentration was 200 $\mu$M, (2) the samples were excited at 370 nm, and (3) emission above 460 nm was detected using a cutoff filter supplied by Durrum.

*Data Fitting.* Using a nonlinear least-squares program NLIN (SAS Institute Inc., Cary, NC), kinetic refolding data at or below 3.0 M urea were fit to a sum of three exponential terms and a constant term reflecting the steady-state signal amplitude:

$$Y(t) = \sum Y_i(0) \exp(-t/\tau_i) + Y_\infty \quad (1)$$

where $Y(t)$ is the total amplitude at time $t$, $Y_\infty$ is the amplitude at infinite time, $Y_i(0)$ is the amplitude corresponding to the individual phase, $i$, at zero time, and $\tau_i$ is the associated relaxation time. Refolding data above 3.0 M urea were fit to a sum of three exponential terms of the type

$$Y(t) = \sum Y_i(0)[\exp(-t/\tau_i)/[1 + q'_i Y_i(0)[1 - \exp(-t/\tau_i)]]] + Y_\infty \quad (2)$$

where the $q'_i Y_i(0)[1 - \exp(-t/\tau_i)]$ term reflects the contribution of the unfolding reaction in the unfolding transition zone (Gittelman & Matthews, 1990; Bernasconi, 1976). The amplitudes of both the CD and fluorescence signals at $t = 0$

were calculated using the following equation:

$$Y_0 = Y_\infty + Y_1(0) + Y_2(0) + Y_3(0) \quad (3)$$

where $Y_0$ is the amplitude at $t = 0$, $Y_\infty$ is the amplitude at infinite time, and $Y_1(0)$, $Y_2(0)$, and $Y_3(0)$ are the calculated amplitude values for each of the three refolding phases observed.

Equilibrium data were fit to a two-state model (Gittelman & Matthews, 1990):

$$D \overset{K}{\rightleftharpoons} 2U$$

where D is the native dimer, U is the unfolded monomer, and $K = [U]^2/[D]$. It can be shown that

$$F_{app} = [(K^2 + 8K[P_{tot}])^{1/2} - K)/(4[P_{tot}]) \quad (4)$$

where $[P_{tot}]$ is the total protein concentration in terms of the monomer. The equations $K = \exp(-\Delta G°_{app}/RT)$ and $\Delta G°_{app} = \Delta G°^{H_2O} + M_G[\text{urea}]$ (Schellman, 1978; Pace, 1986) were substituted into the above equation to calculate $\Delta G°^{H_2O}$ and $M_G$ values. $\Delta G°_{app}$ is the apparent free energy difference at any given urea concentration, $\Delta G°^{H_2O}$ is the free energy difference in the absence of denaturant, and $M_G$ is a parameter that describes the cooperativity of the unfolding transition (Matthews, 1987).

For comparison of the data, the observed changes in circular dichroism and fluorescence signals were converted to an apparent fraction of unfolded protein, $F_{app}$, such that

$$F_{app} = (Y_{obs} - Y_N)/(Y_U - Y_N) \quad (5)$$

where $Y_{obs}$ refers to the observed optical value at a particular urea concentration, and $Y_N$ and $Y_U$ refer to the calculated values for the native and unfolded forms, respectively, at the same denaturant concentration.

To obtain a single comprehensive fitting equation, eqs 4 and 5 above were combined with equations that assume a linear dependence of $Y_N$ and $Y_U$ on the denaturant concentration:

$$Y_N = Y_N^0 + M_N[\text{urea}] \quad (6a)$$

$$Y_U = Y_U^0 + M_U[\text{urea}] \quad (6b)$$

where $M_N$ and $M_U$ are the slopes of the pre- and posttransition regions for the native and unfolded forms, respectively (Santoro & Bolen, 1988). Equation 7 (Reece et al., 1991), which is appropriate for dimeric proteins, shows the dependence of the observed optical parameter on the denaturant concentration:

$$Y_{obs} = [\{[(Y_U^0 + M_U[\text{urea}]) - (Y_N^0 + M_N[\text{urea}])] \times [\{\exp[-(\Delta G°^{H_2O} + M_G[\text{urea}])/RT\}^2 + 8\{\exp[-(\Delta G°^{H_2O} + M_G[\text{urea}])/RT\}[P_{tot}]]^{1/2} - \{\exp[-(\Delta G°^{H_2O} + M_G[\text{urea}])/RT]\}\}/4[P_{tot}]] + Y_N^0 + M_N[\text{urea}] \quad (7)$$

and was used to fit the observed data with a nonlinear least-squares program NLIN (SAS Institute, Inc., Cary, NC). An equation of this same form:

$$Y_{obs} = (Y_U^0 + M_U[\text{urea}]) + [\exp(-(\Delta G°^{H_2O} + M_G[\text{urea}])/RT)/(1 + \exp(-(\Delta G°^{H_2O} + M_G[\text{urea}])/RT))][(Y_U^0 + M_U[\text{urea}]) - (Y_N^0 + M_N[\text{urea}])] \quad (8)$$

was used to fit the monomeric equilibrium data sets. The midpoint of the unfolding transition curves, $C_m$, was obtained by solving for the urea concentration when $F_{app} = 0.5$.

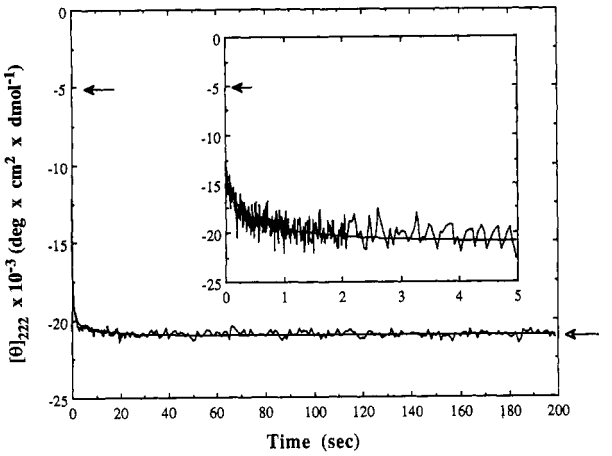



FIGURE 2: Refolding of TR monitored by the ellipticity change at 222 nm as a function of time. Protein was unfolded in 7.0 M urea and subsequently diluted to 2.0 M urea to induce refolding at pH 7.6 and 25 °C (final protein concentration was 0.13 mg/mL). The inset shows kinetic data for the first 5 s after refolding was initiated. The solid arrow indicates the ellipticity value of the unfolded protein in 2.0 M urea and was obtained by linear extrapolation of data collected above 6.0 M urea on the stopped-flow instrument. Open arrow indicates the ellipticity value of a protein sample equilibrated in 2.0 M urea.

## RESULTS

Early events in the folding of globular proteins can often be detected with stopped-flow drive trains integrated into circulated dichroism and fluorescence spectrometers. SF-CD spectroscopy (Labhardt, 1986; Kuwajima et al., 1987) has been used to study the formation of secondary structure in early folding intermediates. A variety of single subunit globular proteins exhibit substantial far UV ellipticity within the dead time of mixing (Sugawara et al., 1991), demonstrating that secondary structure can appear within milliseconds. In the case of $\alpha$-lactalbumin, this burst phase species has been associated with a molten globule-type structure (Kuwajima, 1989).

Another approach to probe the early events in folding involves the use of a hydrophobic dye, ANS, to monitor the formation of hydrophobic sites in proteins [for a review, see Semisotnov et al. (1991)] and in lipid membranes [for a review, see Slavik (1982)]. This application depends upon an increase in the fluorescence intensity and a blue-shift in the wavelength maximum from 520 to 480 nm when an ANS molecule binds to nonpolar sites. Because many hydrophobic residues which are exposed to solvent in unfolded proteins become buried in the native conformation, ANS binding has the potential to monitor the transient formation of hydrophobic clusters during folding. Some of the effects on ANS fluorescence observed for the folding of a number of proteins have been attributed to the formation of the molten globule intermediate (Goldberg et al., 1990; Ptitsyn et al., 1990; Semisotnov et al., 1991).

### *CD Studies*

A representative example of a kinetic trace obtained by measuring the time dependence of the ellipticity at 222 nm during the refolding of TR is shown in Figure 2. The good agreement between the ellipticity observed after the refolded protein achieves a steady-state value and the ellipticity for TR which has not been previously unfolded (Figure 2) demonstrates that the unfolding transition is fully reversible.

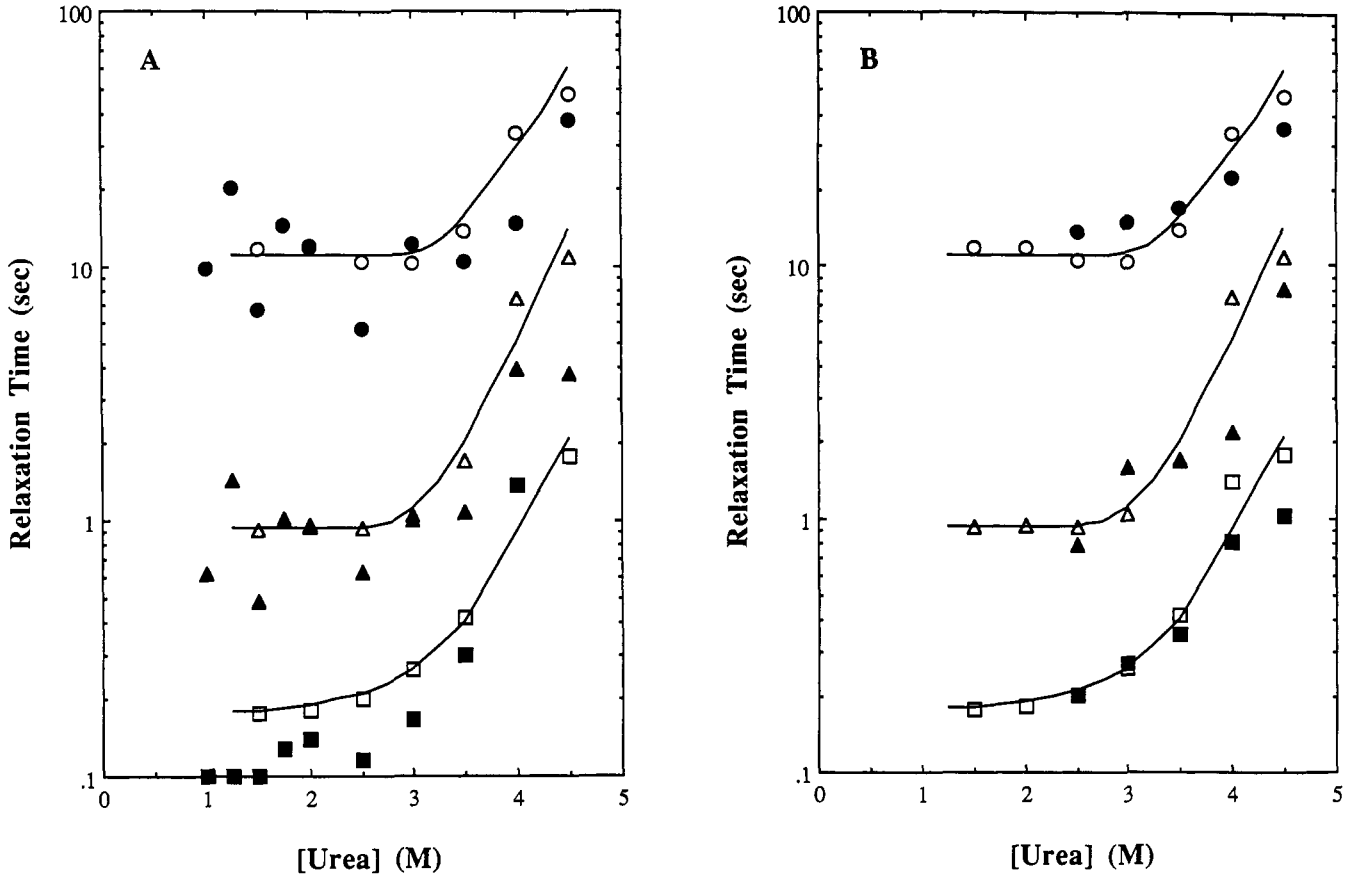



FIGURE 3: (A) Urea dependence of the SF-CD (closed symbols) and SF-Trp Fl (open symbols) refolding relaxation times $\tau_1$ (●, ○), $\tau_2$ (▲, △), and $\tau_3$ (■, □) at pH 7.6 and 25 °C. The final protein concentration was 0.13 mg/mL. Lines through the SF-Trp Fl data are drawn to aid the eye. (B) Urea dependence of the ANS binding (closed symbols) and SF-Trp Fl (open symbols) refolding relaxation times $\tau_1$ (●, ○), $\tau_2$ (▲, △), and $\tau_3$ (■, □) at pH 7.6 and 25 °C. The final protein concentration was 0.13 mg/mL. Lines through the SF-Trp Fl data are drawn to aid the eye.

Depending on the final urea concentration, this and other refolding curves were fit to eq 1 (at and below 3 M urea) or eq 2 (above 3 M urea), using three exponential terms. The ellipticity values extrapolated to zero time, $[\theta]_0$, are more negative than the value of the unfolded protein extrapolated to the same final conditions (2.0 M urea). This discrepancy indicates a significant development of signal within the dead time of mixing for these experiments, 5 ms. The properties of this burst phase species will be described below.

*Refolding Phases Detected by SF-CD.* The relaxation times of the three slower refolding phases following the CD burst phase were plotted as a function of final urea concentration and are shown in Figure 3A. Below 3.0 M urea, all three phases are urea independent, while above 3.0 M urea all three phases exhibit urea dependence. For comparison with these data, SF-Trp Fl. experiments were conducted at the same protein concentration (5.3 $\mu$M). Relaxation times similar to those obtained from SF-CD, including the nonuniform urea dependence, were obtained (Figure 3A). The amplitudes obtained for all three refolding phases detected by SF-CD (data not shown) follow urea-dependent trends similar to those detected by SF-Trp Fl. (Gittleman & Matthews, 1990).

*ANS Studies*

A representation of the ANS fluorescence trace during TR refolding is shown in Figure 4. The fluorescence shows a dramatic increase in intensity over that of free ANS within the dead time of mixing (4 ms), a phenomenon which will be referred to as the ANS burst phase. The fluorescence intensity then increases exponentially and reaches a steady-state value

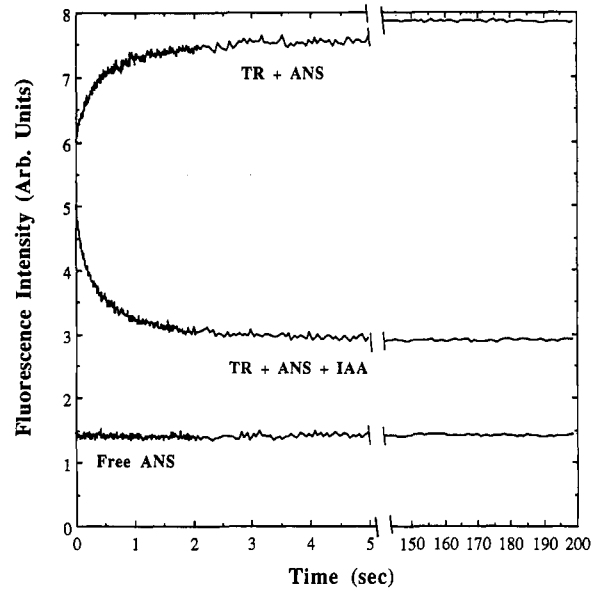



FIGURE 4: Kinetic refolding of TR measured by the change in fluorescence intensity of ANS above 460 nm as a function of time. Protein was unfolded in 7.0 M urea and subsequently diluted to 2.5 M urea to induce refolding in the presence of 200 $\mu$M ANS at pH 7.6 and 25 °C. A kinetic trace is also shown for the refolding of TR in the presence of 200 $\mu$M ANS and 200 $\mu$M IAA. The final protein concentration was 0.13 mg/mL. The intensity of ANS in the same final solution is also shown for reference.

which is much higher than that of the free ANS. This phenomenon very likely reflects the binding of ANS to the

Table I: Kinetic Refolding of TR at 2.5 M Urea

| protein structure formation | technique | relaxation time interval | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | <0.010 s | $A_{BP}$% | 0–0.3 s | $A_3$% | 0.4–2 s | $A_2$% | 3–20 s | $A_1$% |
| secondary structure | far-UV CD (222 nm) | <0.005 | 63 | 0.1 | 16 | 0.6 | 15 | 6 | 6 |
| nonpolar surface | ANS binding (–IAA) | <0.004 | 76 | 0.2 | 9 | 0.8 | 9 | 14 | 6 |
| | ANS binding (+IAA)[a] | <0.004 | 66 | 0.3 | 24 | 1.6 | 10 | nd[b] | nd |
| tertiary structure | Trp fluorescence | nd | nd | 0.2 | 45 | 0.9 | 40 | 11 | 15 |

[a] Absolute values of amplitudes were used to normalize the data for percentage calculations. [b] Not detected.

corepressor L-tryptophan binding site in the native conformation (Chou et al., 1989).

To test this hypothesis, the tryptophan analog indole acrylic acid (IAA) was added along with ANS in the refolding cocktail during the TR refolding process. IAA binds 100 times more tightly ($K_d$ = 0.5 $\mu$M) (Marmorstein et al., 1987) than ANS ($K_d$ = 50 $\mu$M) (Chou et al., 1989) to the corepressor site and will virtually exclude ANS binding at equimolar concentrations. Figure 4 also shows a representative trace of TR refolded in the presence of 200 $\mu$M ANS and 200 $\mu$M IAA. Once again, a significant increase in ANS fluorescence is observed as a burst phase; however, the subsequent fluorescence signal *decreases* exponentially to an intensity slightly higher than free ANS at infinite time. This difference reflects nonspecific binding of ANS molecules to sites on TR other than the cofactor site in the presence of IAA (data not shown). Therefore, the burst phase amplitude in ANS binding primarily, if not entirely, reflects nonspecific binding of the dye to a folding intermediate and not to the corepressor site in the native conformation.

*Refolding Phases Detected by ANS Binding.* The relaxation times of the three slower refolding phases observed following the ANS burst phase were plotted as a function of final urea concentration and are shown in Figure 3B. As was the case for SF-CD spectroscopy, the magnitudes of the relaxation times obtained from fitting ANS data (conducted at the same protein concentration, 5.3 $\mu$M) are similar to those obtained by SF-Trp Fl. In addition, the amplitudes obtained for all three refolding phases detected by ANS binding (data not shown) follow similar urea-dependent trends detected by SF-Trp Fl (Gittelman & Matthews, 1990). The observed relaxation times did not depend on the ANS concentration, indicating that the changes in the ANS fluorescence reflected protein folding reactions and not the binding of ANS to TR (data not shown). Kinetic artifacts encountered below 2.5 M urea for times longer than 5–10 s in the ANS experiments precluded analysis of the three slow refolding phases at lower denaturant concentrations.

Table I summarizes the relative amplitudes and relaxation times obtained from fitting the SF-CD, ANS binding, and SF-Trp Fl refolding data collected at 2.5 M urea. With the possible exception of the binding of ANS in the presence of IAA, good agreement was found between the relaxation times for the slow folding phases detected by these three techniques. Substantial burst phase amplitudes were detected for both SF-CD at 222 nm and for nonspecific ANS binding to nonpolar surfaces. Several unsuccessful attempts were made to detect a burst phase by SF-Trp Fl to probe the involvement of tryptophan interactions in this intermediate. The failure to detect an intrinsic tryptophan fluorescence burst phase was most likely due to inefficient mixing of unfolded protein with high denaturant concentrations required to establish the denatured baseline. The possible exceptions to the agreement between CD, ANS, and intrinsic tryptophan fluorescence data are the slightly longer relaxation times for $\tau_3$ and $\tau_2$ reactions and the absence of a third refolding phase detected by ANS binding in the presence of IAA. Both of these differences may reflect a small fraction of unfolded protein flowing through the $\tau_1$ channel (15% as measured by intrinsic tryptophan fluorescence), making this reaction difficult to detect.

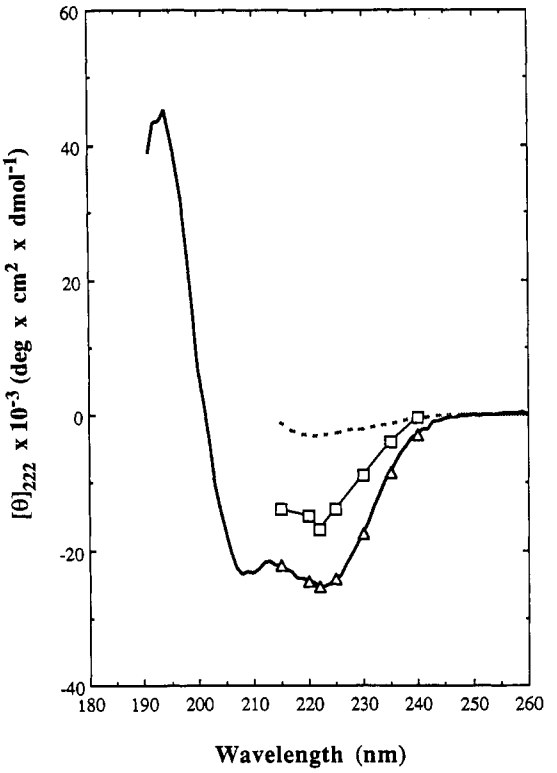



FIGURE 5: Peptide CD spectra of TR at pH 7.6 and 25 °C. Native state (—) and unfolded state in 7 M urea (- - -). The burst phase amplitudes, $[\theta]_0$ (□) and the final ellipticities, $[\theta]_\infty$ (△), in refolding to 2.0 M urea are also shown. The final protein concentration was 0.13 mg/mL.

*Characterization of the Burst Phase Species*

Over half of the total signal change observed in SF-CD and ANS binding experiments occurred within the dead time of mixing (4–5 ms), indicating the formation of a significant population of a partially folded intermediate within this time range. Because the burst phase intermediate forms at least two orders of magnitude more rapidly than any subsequent steps in folding, the amplitude of this phase can be taken as a measure of the concentration of this species. The dependence of this amplitude on the final urea concentration, therefore, provides a measure of its stability (Kuwajima et al., 1985; Ikeguchi et al., 1986).

*Burst Phase Species Monitored by SF-CD*

*Secondary Structure in Burst Phase Species.* The CD spectra of TR in the native and fully unfolded conformations are shown in Figure 5. The native spectrum exhibits the double minima at 208 and 222 nm, typical of highly $\alpha$-helical proteins

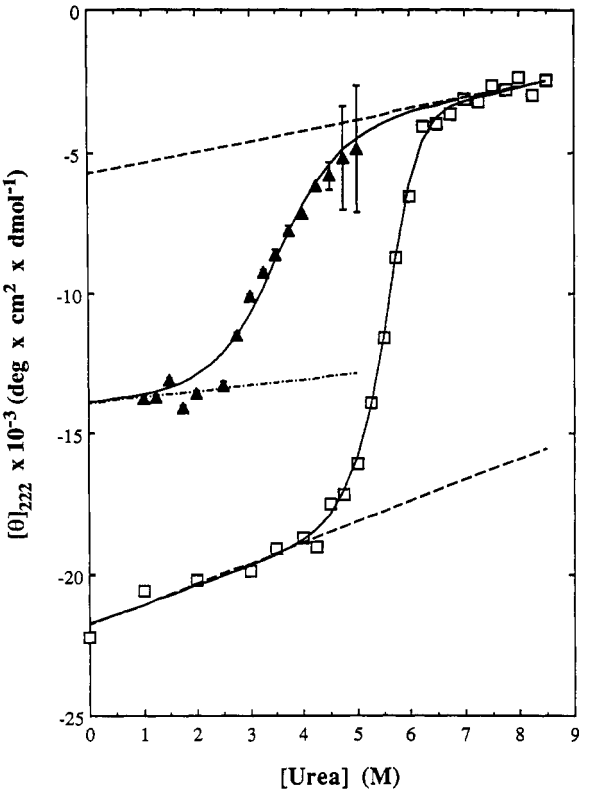



FIGURE 6: Equilibrium unfolding transition of TR (□) monitored by the change in the ellipticity at 222 nm as a function of urea concentration at pH 7.6 and 25 °C. The final protein concentration was 0.13 mg/mL. The burst phase amplitudes $[\theta]_0$ observed in refolding to a final protein concentration of 0.13 mg/mL (▲) are also indicated. Error bars indicate the fitting errors associated with the burst phase amplitude calculations. Solid lines correspond to the fit of the data sets, as described under Materials and Methods. Dashed lines indicate the native and unfolded baselines used for data fitting; the dotted line indicates the folded baseline used for the data fitting of the burst phase intermediate.

and peptides (Greenfield & Fasman, 1969). In the presence of 7.0 M urea, the ellipticity at 222 nm is relatively small ($[\theta]_{222} = -3.1 \times 10^{-3}$ deg·cm$^2$·dmol$^{-1}$), indicating that the protein has little or no secondary structure remaining. The $[\theta]_0$, or burst phase amplitudes, obtained from kinetic data collected at different wavelengths but the same final urea concentration, 2.0 M, are also shown in Figure 5. Even with the limitation that the denaturant prevents detection below 215 nm, a minimum is evident in the 222-nm region. The similarity to the spectrum of native TR indicates the formation of $\alpha$-helices within the dead time of these stopped-flow experiments (5 ms). The amplitude at 2.0 M urea accounts for slightly over half of the signal found either in the native conformation or under equilibrium conditions at the end of the kinetic run (Figure 5).

*Stability of CD Burst Phase Species.* The $[\theta]_0$ values at 222 nm obtained by varying the final urea concentration are shown in Figure 6. Between 1.0 and 2.0 M urea, the amplitude of the burst phase CD signal at 222 nm is independent of the denaturant concentration. Above 2.0 M urea, the signal diminishes in a sigmoidal fashion to an undetectable level above 5.0 M urea.

A quantitative estimate of the stability of this species requires knowledge of the order of the reaction, i.e., whether this intermediate is monomeric or dimeric. Although previous studies of TR have shown that the development of the dimer is associated with the slower, rate-limiting reactions described above (Gittelman & Matthews, 1990), the order was directly probed by examining the dependence of this transition curve on the protein concentration. Virtually identical transition curves were obtained at 0.06, 0.13, and 0.26 mg/mL TR (data not shown), supporting the conclusion that this burst phase intermediate is a monomer. Data from quench-flow refolding experiments involving wild-type TR and a mutant containing an additional positive charge, G85R, are also consistent with a monomeric species. These two proteins were independently unfolded, subjected to refolding for varying periods of time, and then mixed together to complete the folding reaction. Separating the two homodimers and the heterodimers by gel electrophoresis, it was found that heterodimers are capable of forming up until the point where the delay is sufficiently long to allow the slow folding reactions to occur; only homodimers are detected when the slow folding reactions precede the mixing step (C. J. Mann, and C. R. Matthews, unpublished results).

Assuming a two-state model of the form

$$M_{BP} \overset{K}{\rightleftharpoons} U \qquad (11)$$

where $M_{BP}$ is the monomeric burst phase intermediate observed within the dead time of mixing, U is the unfolded monomer, and $K = [U]/[M_{BP}]$, the pre- and posttransition baselines, as well as the transition itself, were simultaneously fit to the appropriate equation (see Materials and Methods; Santoro & Bolen, 1988) using a nonlinear least-squares program. Note that the posttransition baseline was taken to be that observed for equilibrium measurements (Figure 6).

At pH 7.6 and 25 °C, and assuming a linear dependence of the free energy of folding on the denaturant concentration (Schellman, 1978), the stability of the burst phase species (in the absence of denaturant) was found to be 3.6 ± 0.3 kcal/mol. The midpoint of the unfolding transition occurs at 3.6 M urea, and the cooperativity parameter, $A$, is 1.0 ± 0.1 kcal/mol·M (urea). For comparison, the equilibrium unfolding transition of native, dimeric TR was also followed by measuring the mean residue ellipticity at 222 nm ($[\theta]_{222}$), as a function of the final urea concentration (Figure 6). Under the same conditions, the midpoint of the unfolding transition occurs at 5.6 M urea, $\Delta G^{\circ}{}_{H_2O}$ = 22.4 ± 0.7 kcal/mol, and $A$ = 2.8 ± 0.1 kcal/mol·M (urea). These results agree well with those previously reported for the native TR dimer (Gittelman & Matthews, 1990). Thus, the burst phase intermediate has approximately 15% of the stability of the native form.

### *Burst Phase Species Monitored by ANS Binding*

*Stability of ANS Burst Phase Species.* The dependence of the amplitude of the ANS burst phase on the final urea concentration is shown in Figure 7A. The amplitude decreases in a monotonic fashion between 1.7 and 5 M urea, indicating a loss in binding of the dye to nonpolar surfaces. The nature of this dependence, with an inflection at 2 M urea and an approach to a limiting value above 5 M urea, suggests that two phenomenon are occurring. Below 2 M urea, the decrease in ANS fluorescence reflects the washing of the ANS off of the nonpolar surfaces which are exposed in the intermediate. The enhanced solubility of such surfaces and of the hydrophobic dye in increasing concentrations of urea would account for this result. The subsequent decrease in amplitude above 2 M urea then reflects the actual unfolding of the intermediate and the loss of nonpolar surfaces which are sufficient in size to bind the dye. Semisotnov et al. (1991) have previously shown that unfolded proteins are not capable of binding ANS. The linear dependence of the fluorescence intensity of the dye

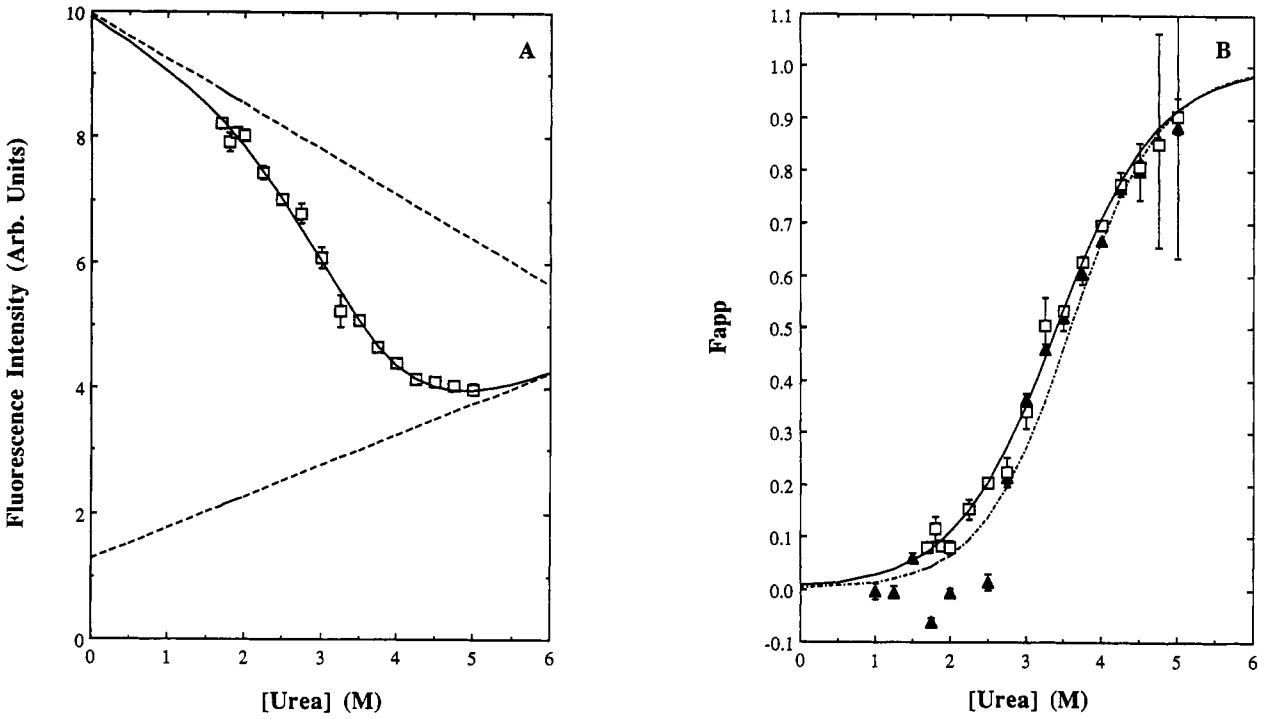



FIGURE 7: (A) Dependence of the ANS burst phase fluorescence intensity on urea concentration at pH 7.6 and 25 °C. The burst phase amplitudes from 1.7 to 2.5 M urea were obtained by fitting the data up to the first 5 s to a pair of exponentials corresponding to the fast and intermediate folding phases and extrapolated to zero time. Solid line indicates the fit of the data as described under Materials and Methods. Dashed lines indicate the folded and unfolded baselines used for data fitting. (B) Dependence of the apparent fraction unfolded protein, $F_{app}$, of the burst phase intermediate on the urea concentration as monitored by ANS binding (□) and SF-CD (▲). Solid line indicates the fit of the ANS data and dotted line indicates the fit of the SF-CD data to a two-state model as described under Materials and Methods.

above 5 M urea approaches that expected for free ANS in urea/water solutions (data not shown), demonstrating that the intermediate is completely unfolded under these condition.

Making these assumptions, these data were fit to a two-state model of the form

$$M_{BP} \overset{K}{\rightleftharpoons} U$$

where $M_{BP}$ is the monomeric burst phase intermediate observed within the dead time of mixing, U is the unfolded monomer, and $K = [U]/[M_{BP}]$. Assuming that the unfolded baseline is similar to that for free ANS, the pre- and posttransition baselines as well as the transition region were simultaneously fit (Santoro & Bolen, 1988) with a nonlinear least-squares program. Note that limitations in the solubility of the TR/ANS complex preclude the collection of data below 1.7 M urea.

At pH 7.6 and 25 °C, and assuming a linear dependence of the free energy of folding on the denaturant concentration (Schellman, 1978), the stability of the burst phase species (in the absence of denaturant) was found to be 3.0 ± 0.1 kcal/mol. The midpoint of the unfolding transition occurs at 3.4 M urea, and the cooperativity parameter, $A$, is 0.9 ± 0.1 kcal/mol·M (urea). A graphical comparison of the ANS and CD data was made by normalizing the data to the apparent fraction of unfolded protein, $F_{app}$ (Gittelman & Matthews, 1990), and plotting $F_{app}$ as a function of final urea concentration (Figure 7B). The $F_{app}$ curves, similar to the above fitting parameters, are in good agreement with each other. The small deviations of the CD burst phase values at 1.75, 2.0, and 2.5 M urea from the fitted transition curve (Figure 6) and subsequent normalized $F_{app}$ curve (Figure 7B) reflect the reproducibility error. Within the experimental error, the secondary structure and the nonpolar surface in the burst phase species are disrupted in a concerted, two-state fashion.

## DISCUSSION

Stopped-flow far-UV CD and ANS fluorescence studies of the folding of TR reveal the presence of a stable, monomeric intermediate which appears within the first 5 ms of the folding reaction. The coincidence of the burst phase unfolding transition curves for CD and ANS fluorescence has two important implications: (1) the process can be accurately described by a two-state model, and (2) the secondary and nonspecific hydrophobic tertiary structure in this intermediate are simultaneously disrupted by urea.

The first observation implies that this folding intermediate is a well-defined thermodynamic state and that a quantitative assessment of its stability can be made. Fernando and Royer (1992) have previously found that dilution of TR results in tryptophan fluorescence changes which are consistent with the dissociation of the native dimer into monomers. The upper limit of the dimer dissociation constant is estimated to be 1 nM, implying a free energy of association of 12.4 kcal/mol. Given that the stability of TR under standard-state conditions is 22.4 kcal/mol, the discrepancy in these energies implies that the dissociated monomers remain folded and have a stability of 5.0 kcal/mol monomer. This value is greater than that for the stability of the burst phase species and suggests that monomers may undergo further folding before the association reaction.

The concerted loss of higher order structure is consistent with an important role for a species in which helical structure and nonpolar surfaces are tightly coupled. One possibility is a hydrophobic core upon which amphipathic helices are stabilized through nonspecific association of their nonpolar side chains. Such a species would have some of the characteristics of a molten globule (Kuwajima, 1989). In addition to the high secondary structure content and nonpolar surfaces, the value of the cooperativity parameter, $A$, is approximately

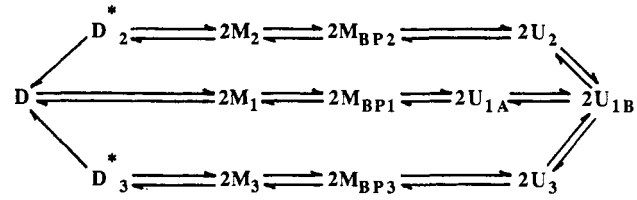
one-third that for the unfolding of the native dimer (Figure 6). If this parameter is a measure of the change in exposure of hydrophobic surface to solvent upon unfolding (Meeker & Shortle, 1986; Kuwajima et al., 1989), then a substantial fraction of the hydrophobic surfaces in the burst phase intermediate, $\sim^1/_3$, must be buried. Another possible explanation for the CD signal is a species of such a high degree of compaction that secondary structures are favored by optimal packing requirements (Chan & Dill, 1990). The latter explanation seems less likely because packing densities near that of the native conformation appear to be required (Chan & Dill, 1990; Gregoret & Cohen, 1991). Considering the highly intertwined nature of the subunits in the native conformation (Zhang et al., 1987), it is difficult to imagine that the packing in a collapsed monomer would approach that found in the native form.

SF-CD has been used extensively by Kuwajima and coworkers to study the rapid formation of secondary structure in a number of proteins, including $\alpha$-lactalbumin (Kuwajima et al., 1985; Ikeguchi et al., 1986), lysozyme (Kuwajima et al., 1985; Ikeguchi et al., 1986), tryptophan synthase $\beta_2$ subunit (Goldberg et al., 1990), staphylococcal nuclease A (Sugawara et al., 1991), and most recently, dihydrofolate reductase (Kuwajima et al., 1991). For each of these proteins, a burst phase intermediate was detected within the experimental dead time. Sigmoidal transition curves for $[\theta]_0$ were observed for $\alpha$-lactalbumin and lysozyme (Ikeguchi et al., 1986), but not for parvalbumin (Kuwajima et al., 1988) or dihydrofolate reductase (Kuwajima et al., 1991). Because these early intermediates are estimated to have only marginal stability even when sigmoidal curves are observed, it is not clear whether this difference is simply a reflection of sequence variations or more of a fundamental phenomenon.

A recent study of the folding of apomyoglobin using a combination of SF-CD and quench-flow hydrogen-exchange NMR techniques (P. A. Jennings and P. E. Wright, unpublished results) has also demonstrated the presence of a stable, burst phase intermediate with substantial secondary structure. Protection of amide protons in the sequences corresponding to helices A, G, and H in the native conformation against exchange shows that the secondary structure is native-like. In related experiments, Hughson et al. (1991) have shown that the A, G, and H helices formed in the stable acid intermediate of apomyoglobin appear to be stabilized by loose hydrophobic interactions, rather than by the native-like helix packing interactions first proposed (Hughson et al., 1990). Assuming that these two intermediates are in fact identical, it would be very interesting to learn if the same types of loose hydrophobic interactions exist in the TR burst phase species. A combination of mutagenesis and quench-flow NMR methods may provide the answer.

Although ANS binding has been used in the past to highlight the existence of nonpolar surfaces in folding intermediates (Ptitsyn et al., 1990; Tandon & Horowitz, 1989), no attempt has been made to use the enhanced fluorescence to measure the stability of a partially folded protein. The principal concern with its accuracy is whether the binding required to obtain the signal contributes to the apparent stability. An upper estimate on the effect can be obtained by measuring a binding isotherm for ANS in a refolding experiment. The amplitudes of all of the ANS phases changed in proportion to the relative concentration of ANS/TR and followed simple binding isotherms. The systematic variation of the ANS concentration between 50 and 300 $\mu$M at constant protein concentration (5.3 $\mu$M) and 2.5 M urea, demonstrated that the dissociation constant must be greater than 300 $\mu$M. Assuming that a single ANS molecule binds preferentially to the burst phase intermediate, the increase in the apparent stability would be less than 0.3 kcal/mol at 200 $\mu$M ANS. Because this value is of the same magnitude as the experimental error in the estimated stability, it can be concluded that the ANS transition curve is an accurate reflection of the thermodynamic properties of the burst phase species.

*Kinetic Folding Mechanism of TR.* The SF-CD and ANS binding data reported in the present paper can be incorporated along with prior results (Gittelman & Matthews, 1990) into a working model for the folding reaction:




where $U_{1-3}$ denote a set of monomeric unfolded forms, $M_{BP1-3}$ a set of monomeric burst phase intermediates detected by SF-CD and ANS binding, and $M_{1-3}$ a set of fully folded monomers; $D^*_2$ and $D^*_3$ are dimer intermediates, and D is the native dimer. All three of the dimeric species are proposed to bind ANS in the corepressor site and, therefore, be native-like in this region.

The unfolded forms are postulated to be kinetically distinct species which fold more rapidly than they can interconvert. Proline isomerization (Brandts et al., 1975) provides one possible explanation; however, this has not been directly demonstrated for TR. Within 5 ms, three burst phase intermediates appear which are subsequently processed to three fully folded monomeric forms. These latter species are presumed to exist, on the basis of the discrepancy between the stability of the burst phase species and the estimated stability of the fully folded monomer (see above). No assignable spectroscopic signal has been detected for this transition. The folded monomers in the middle channel associate to form the native dimer while those in the upper and lower channels form dimeric intermediates, $D^*_2$ and $D^*_3$. These latter species then interconvert to the native dimer.

The detection of three virtually identical slow folding phases by far-UV CD, Trp fluorescence, and ANS binding to the corepressor site shows that similar events are occurring in each channel. Because these techniques are sensitive to the formation of secondary and tertiary structure during folding, it also appears that both classes of structural changes occur in a concerted fashion for each channel. The observation that the relaxation times are in close agreement in the range of urea concentrations where the dimerization reaction is limiting (above 3 M urea; Gittelman & Matthews, 1990) implies that the formation of quaternary structure is also coordinated with that of the final, native secondary and tertiary structure.

Detection of a stable, monomeric burst phase intermediate in TR by SF-CD and ANS binding techniques demonstrates that folding does indeed occur prior to subunit association. This species appears within 4–5 ms, contains at least half of the secondary structure of native dimer, and has buried a significant portion ($\sim^1/_3$) of its nonpolar surface. The concerted disruption of secondary and tertiary structure by urea implies that this intermediate is a distinct thermodynamic state with a free energy of $\sim$3.3 kcal/mol. Independent estimates of the dimer/monomer dissociation constant in the absence of denaturant (Fernando & Royer, 1992) suggest

that a fully folded monomer develops an additional ~1.7 kcal/mol of stability prior to the dimerization event. The conformational changes implied by this putative step in folding have thus far been spectroscopically undetectable. The remainder of the free energy of folding, 12.4 kcal/mol, appears in the rate-limiting dimerization reaction. These data suggest that the incremental and progressive decrease in free energy of monomeric intermediates may be an important factor in the rapid and efficient folding of oligomeric protein systems.

## ACKNOWLEDGMENT

We thank Drs. Mitch Gittelman, Cathy Royer, and Pat Jennings for helpful discussions. We also thank Pat Jennings for sharing her unpublished results and Bryan Jones and the rest of the Matthews group for their helpful insight.

## REFERENCES

Bernasconi, C. F. (1976) *Relaxation Kinetics*, Academic Press, New York.
Blond, S., & Goldberg, M. E. (1985) *J. Mol. Biol. 182*, 587–606.
Brandts, J. F., Halvorson, H. R., & Brennan, M. (1975) *Biochemistry 14*, 4953–4963.
Chan, H. S., & Dill, K. A. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 6388–6392.
Chou, W.-Y., Bieber, C., & Matthews, K. S. (1989) *J. Biol. Chem. 264*, 18309–18313.
Fernando, T., & Royer, C. (1992) *Biochemistry 31*, 3429–3441.
Freemont, P., Lane, A. N., & Sanderson, M. (1991) *Biochem. J. 278*, 1–24.
Gittelman, M. S., & Matthews, C. R. (1990) *Biochemistry 29*, 7011–7020.
Goldberg, M. E., Semisotnov, G. V., Friguet, B., Kuwajima, K., Ptitsyn, O. B., & Sugai, S. (1990) *FEBS Lett. 263*, 51–56.
Greenfield, N. J., & Fasman, G. D. (1969) *Biochemistry 8*, 4108–4116.
Gregoret, L. M., & Cohen, F. E. (1991) *J. Mol. Biol. 219*, 109–122.
Gunsalus, R. P., & Yanofsky, C. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 7717–7721.
Haugland, R. P. (1991) *Handbook of Fluorescent Probes and Research Chemicals*, Molecular Probes, Inc., Catalog, Eugene, OR.
Hughson, F. M., Wright, P. E., & Baldwin, R. L. (1990) *Science 249*, 1544–1548.
Hughson, F. M., Barrick, D., & Baldwin, R. L. (1991) *Biochemistry 30*, 4113–4118.
Ikeguchi, M., Kuwajima, K., Mitani, M., & Sugai, S. (1986) *Biochemistry 25*, 6965–6972.
Jaenicke, R. (1987) *Prog. Biophys. Mol. Biol. 49*, 117–237.
Jaenicke, R. (1991) *Biochemistry 30*, 3147–3161.
Joachimiak, A., Kelley, R. L., Gunsalus, R. P., Yanofsky, C., & Sigler, P. B. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 668–672.
Kelley, R. L., & Yanofsky, C. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 483–487.
Kuwajima, K., Sakuraoka, A., Fueki, S., Yoneyama, M., & Sugai, S. (1988) *Biochemistry 27*, 7419–7428.
Kuwajima, K., Hiraoka, Y., Ikeguchi, M., & Sugai, S. (1985) *Biochemistry 24*, 874–881.
Kuwajima, K., Yamaya, H., Miwa, S., Sugai, S., & Nagamura, T. (1987) *FEBS Lett. 221*, 115–118.
Kuwajima, K., Sakuraoka, A., Fueki, S., Yoneyama, M., & Sugai, S. (1988) *Biochemistry 27*, 7419–7428.
Kuwajima, K., Mitani, M., & Sugai, S. (1989) *J. Mol. Biol. 206*, 547–561.
Kuwajima, K., Garvey, E. P., Finn, B. E., Matthews, C. R., & Sugai, S. (1991) *Biochemistry 30*, 7693–7703.
Labhardt, A. M. (1986) *Methods Enzymol. 131*, 126–135.
Le Bras, G., Teschner, W., Deville-Bonne, D., & Garel, J.-R. (1989) *Biochemistry 28*, 6836–6841.
Matthews, C. R. (1987) *Methods Enzymol. 154*, 498–511.
Marmorstein, R. Q., Joachimiak, A., Sprinzl, M., & Sigler, P. B. (1987) *J. Biol. Chem. 262*, 4922–4927.
Meeker, A. K., & Shortle, D. (1986) *Proteins: Struct., Funct., Genet. 1*, 81–89.
Pace, C. N. (1986) *Methods Enzymol. 131*, 266–280.
Paluh, J. L., & Yanofsky, C. (1986) *Nucleic Acids Res. 14*, 7851–7860.
Ptitsyn, O. B., Pain, R. H., Semisotnov, G. V., Zerovnik, E., & Razgulyaev, O. I. (1990) *FEBS Lett. 262*, 20–24.
Reece, L. J., Nichols, R., Ogden, R. C., & Howell, E. E. (1991) *Biochemistry 30*, 10895–10904.
Rudolph, R., Fuchs, I., & Jaenicke, R. (1986) *Biochemistry 25*, 1662–1669.
Santoro, M. M., & Bolen, D. W. (1988) *Biochemistry 27*, 8063–8068.
Schagger, H., & von Jagow, G. (1987) *Anal. Biochem. 166*, 368–379.
Schellman, J. A. (1978) *Biopolymers 17*, 1305–1322.
Semisotnov, G. V., Rodionova, N. A., Razgulyaev, O. I., Oversky, V. N., Gripas, A. F., & Gilmanshin, R. I. (1991) *Biopolymers 31*, 119–128.
Slavik, J. (1982) *Biochim. Biophys. Acta 694*, 1–25.
Steitz, T. A. (1990) *Q. Rev. Biophys. 23*, 205–280.
Sugawara, T., Kuwajima, K., & Sugai, S. (1991) *Biochemistry 30*, 2698–2706.
Tandon, S., & Horowitz, P. M. (1989) *J. Biol. Chem. 264*, 9859–9866.
Tonomura, B., Nakatani, H., Ohnishi, M., Yamaguchi-Ito, J., & Hiromi, K. (1978) *Anal. Biochem. 84*, 370–383.
Zetina, C. R., & Goldberg, M. E. (1980) *J. Mol. Biol. 137*, 401–414.
Zetina, C. R., & Goldberg, M. E. (1982) *J. Mol. Biol. 157*, 133–148.
Zhang, R.-G., Joachimiak, A., Lawson, C. L., Schevitz, R. W., Otwinowski, Z., & Sigler, P. B. (1987) *Nature 327*, 591–597.