Folding Kinetics of Phage T4 Thioredoxin†

K. L. B. Borden and F. M. Richards*
Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut 06511
Received October 6, 1989; Revised Manuscript Received November 27, 1989

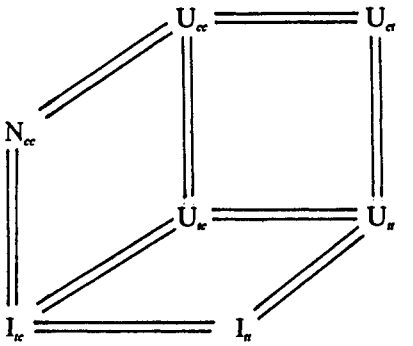
ABSTRACT: The folding mechanism for bacteriophage T4 thioredoxin is best described by a four-state box mechanism, $N \rightarrow U_c \rightarrow U_t \rightarrow I_t \rightarrow N$, where N indicates native, U_c the unfolded form with the cis proline isomer, U_t unfolded with the trans proline isomer, and I_t a compact form with a trans proline isomer. Both manual mixing fluorescence and size-exclusion chromatography indicate that there is a cis–trans proline isomerization that is important to the folding pathway. Furthermore, the data suggest that the cis–trans isomerization can also occur in a compact nativelike state which is refered to as I_t. The slow phase seen in fluorescence seems to be monitoring the cis–trans isomerization in the compact form, not the isomerization which occurs in the denatured state.

The crystal structures of both *Escherichia coli* thioredoxin, EcTrx (Holmgren & Branden, 1975; Katti et al., 1989), and bacteriophage T4 thioredoxin, T4Trx (Söderberg & Branden, 1978), are known. The secondary structure of residues 22–108 of EcTrx is similar to that of residues 1–87 of T4Trx which is missing the first 21 residues of the *E. coli* protein (Holmgren et al., 1975; Söderberg et al., 1978). In this 87-residue region of structural homology, there are only 9 residues in common between the 2 proteins. However, like its *E. coli* counterpart, T4Trx has a single disulfide group in a short peptide loop which is central to its redox function. LeMaster (1986) has developed a strain of *E. coli* and appropriate plasmid, pDL59, for the simultaneous overproduction of EcTrx and T4Trx. This system provides an opportunity to study how proteins with very different amino acid sequences arrive at similar tertiary structures.

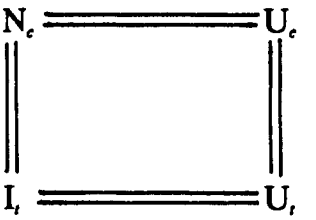
Stellwagen and co-workers have extensively studied the folding pathway of EcTrx using both fluorescence and size-exclusion chromatographic techniques (Kelley & Stellwagen, 1984; Kelley et al., 1986; Shalongo et al., 1987). They have shown that the folding data are best fit by the "cubic" reaction cycle shown in Scheme I (Shalongo et al., 1987). N and U denote native and denatured species, respectively. I is a compact intermediate containing one or more nonnative proline isomers. The subscripts c and t denote the cis or trans forms of two proline residues presumed to undergo isomerization. Of the four proline residues in the protein, the isomerization of Pro-76 has been shown to be responsible for the slowest step in the folding pathway (Kelley & Richards, 1987). A second proline residue is thought also to be involved but has not yet been positively identified.

The two proteins have different spectral properties as expected from the amino acid composition. EcTrx has two tryptophan and two tyrosine residues while T4Trx has four tyrosines and no tryptophans. These aromatic residues are not in analogous positions in the 3D structure of the two proteins (see Figure 9). The fluorescence changes seen in EcTrx are thought to be due to the two trp residues while those in T4Trx can only reflect changes in the environment of the Tyr residues. In spite of the substantially lower magnitude, the fluorescence changes in T4Trx can be used to monitor the

Scheme I



Scheme II



folding process. This paper reports both fluorescence and size-exclusion chromatographic studies of the folding of T4Trx. We will conclude that the present data are adequately described by the reactions shown in Scheme II. The symbol I represents an intermediate with nativelike volume but with a trans isomer of proline-66 which is cis in the native protein.

Materials and Methods

LeMaster (1986) has described the isolation and cloning of the bacteriophage T4 thioredoxin gene nrdC. This gene along with the *E. coli* thioredoxin gene, trxA, was placed in the plasmid pDL59 (LeMaster & Richards, 1988). The bacterial growth and protein purification were carried out according to LeMaster and Richards (1988).

All protein samples were made up in 50 mM sodium phosphate buffer, pH 6.0. Except in multimixing experiments, samples were equilibrated overnight in the appropriate guanidine hydrochloride (GuHCl) concentration, and for renaturation were diluted and allowed to stand overnight again before measurement. Guanidine hydrochloride was ultrapure grade from Sigma. GuHCl concentration was determined by refractive index measurements.

Fluorescence. Equilibrium fluorescence measurements were conducted on a Perkin-Elmer MPF3 spectrophotometer with

† This work was supported by National Institute of General Medical Sciences Grant GM-22778, F.M.R. principal investigator.
* Correspondence should be addressed to this author at the Department of Molecular Biophysics and Biochemistry, Yale University, 260 Whitney Ave., New Haven, CT 06511.

0006-2960/90/0429-3071$02.50/0

1-cm path-length cells. A circulating water bath held the cell temperature at 25 or 4 °C. At the lower temperature, dry nitrogen was circulated through the sample chamber in order to prevent fogging. The emission wavelength was 304 nm, and the excitation wavelength was 280 nm. Slit widths for emission and excitation were 10 and 4 nm, respectively.

The protein concentration varied from 2 to 15 μM and was measured on a Varian 219 Cary spectrophotometer by monitoring the absorbance at 280 nm. The molar absorptivity was assumed to be 6500 M^{-1} cm^{-1} (Berglund & Söderberg, 1970). Fluorescence intensity from the chart record was normalized by the protein concentration.

From the plots of intensity versus denaturant concentration, the fluorescence of the sample in the native state, y_N, and the denatured state, y_D, was determined from the asymptotic values at low and high concentration, respectively. With the known protein concentration, these yielded the calibration factors h_N and h_D in chart units per micromolar protein for later use in the kinetic experiments. With the assumption of a two-state process under these equilibrium conditions, one can estimate the fractions of the sample in the folded, f_N, and unfolded, f_U, states:

$$f_U = \frac{y - y_N}{y_D - y_N}; \quad f_N = \frac{y_D - y}{y_D - y_N}; \quad f_U + f_N = 1 \qquad (1)$$

and

$$K_{unf} = f_U/f_N \qquad \text{or} \qquad \Delta G_{unf} = -RT \ln (f_U/f_N) \qquad (2)$$

where K_{unf} and ΔG_{unf} are the equilibrium constant or the equivalent free energy change for the unfolding reaction at the particular denaturant concentration of the experiment.

In kinetic measurements using manual mixing, all samples were filtered by using 0.22-μm filters and mixed via inversion. The dead time varied from 20 to 30 s. To interpret the recorded data of fluorescence intensity versus time, the chart traces were digitized by hand. Within experimental error, the plots of the logarithm of fluorescence intensity versus time were linear over the time intervals recorded and were fitted by linear regression to a single exponential of the form $y = A \exp(-t/\tau)$, where τ is the time constant for the reaction. The large range of relaxation times was covered in a series of different experiments, each of which could be fit by a single exponential within the error of the data. The slowest time constants are denoted by the subscript 1 with increasing numbers referring to faster phases. To minimize instrumental drift, the final intensity, f_∞, was estimated from the plateau level after 20 min. The rates at 25 °C were all fast enough that 20 min was effectively infinite time. However, in the 4 °C studies, the final intensity, f_∞, was measured at 2 h.

The initial intensity at $t = 0$, f_0, was defined as the original fluorescence of either the native or the denatured protein depending on whether the experiment was following the unfolding or the folding reaction. The value of f_0 was calculated by multiplying the known protein concentration by the calibration factor h_N or h_D previously determined in the equilibrium experiments.

Following Kelley and Richards (1987), the phase amplitude, α_i, was calculated as

$$\alpha_i = \frac{f_i - f_\infty}{f_0 - f_\infty} \qquad (3)$$

where f_∞ is the y intercept of the semilog plot. Each time constant τ_i has a corresponding α_i determined in the same run.

The stopped-flow kinetic measurements were made on a DG110 rapid-mixing fluorometer at the University of Connecticut, thanks to T. Schuster. The mixing device provided 1:1 mixing ratios. The excitation wavelength used was 280 nm. A Pyrex glass filter between the observation chamber and the photomultiplier tube was used to block the scattered excitation light. All samples were filtered with 0.22-μm filters as usual, and all buffers were degassed for at least half an hour. Protein concentration for these experiments ranged from 4 to 15 μM. Samples were equilibrated in a 20 °C water bath for at least 20 min before mixing to prevent schlieren effects. Spectra were recorded by taking pictures of the oscilloscope traces and worked up as described for the chart records under manual mixing. The fast and slow kinetic phases could not both be seen on a single trace because the time constants were too far apart; an f'_∞ was estimated which was the apparent asymptote to the faster decay (refolding only) and was used in the time constant and amplitude calculations for the fast phase. To determine the final value, f_∞, the scope was triggered 20 min later to ensure that the slower folding phase had gone to completion. (The oscilloscope did not have a time base slow enough to monitor this reaction continuously.)

Circular Dichroism. Measurements for both the near- and far-UV regions were performed on an Aviv 60DS spectropolarimeter at the Yale Instrumentation Center. The spectra were collected at ambient temperature, ranging from 25 to 30 °C. In the far-UV, 1-mm path-length cells were used, and protein concentrations were close to 20 μM. Scans were made from 250 to 200 nm. Due to the absorbance at higher GuHCl concentrations, readings extended only to 210 nm. In the near-UV, 1-cm path-length cells were used with scans from 310 to 260 nm.

Size-Exclusion Chromatography. The procedure is described in detail by Shalongo et al. (1987). The experiments were carried out in the laboratory of E. Stellwagen at the university of Iowa. The separations were done at 4 °C with the column in a cold box. Absorbance was monitored at 225 nm, an isosbestic point. The experiments were arranged in sets. Each set consisted of (1) an unfolding experiment, (2) a refolding experiment, (3) an equilibrium experiment, and (4) a 30-s pulse-denature experiment. In (1), a native protein sample was applied to the column; in (2), a sample of denatured protein; in (3), a sample equilibrated in the chromatographic running buffer; and in (4), a sample placed in 6.2 M GuHCl for 30 s before being loaded on the column. The separation pattern was obtained in about 10 min after sample loading. The native protein elutes at about 8.5 min and the denatured protein at 7.5 min. The elution times are somewhat solvent dependent, but there is typically a 1-min separation between the elution times of the native and denatured forms. A transition width of 1 M or less was used in the simulations.

Results

The equilibrium unfolding transition is shown in Figure 1. The process was monitored by changes in fluorescence and in circular dichroism spectra in both the near- and far-UV. Due to interference from Raman scattering in the fluorescence measurements, shifts in the maximum wavelength of the fluorescence as denaturation proceeds could not be resolved. Another technical problem is the weakness of the tyrosine signal in both the near UV CD and fluorescence. The actual data with error bars are shown in panels A–C of Figure 1, while the derived curves for f_u (eq 1) for all three data sets are shown in panel D.

The fluorescence intensity of T4 thioredoxin increases by 2-fold upon denaturation at 25 °C. Renaturation studies indicate that folding is reversible. The transition midpoint is at 1.2 M GuHCl (Figure 1). The transition is unusually sharp with the entire transition occurring in a range of 0.7–1.6 M

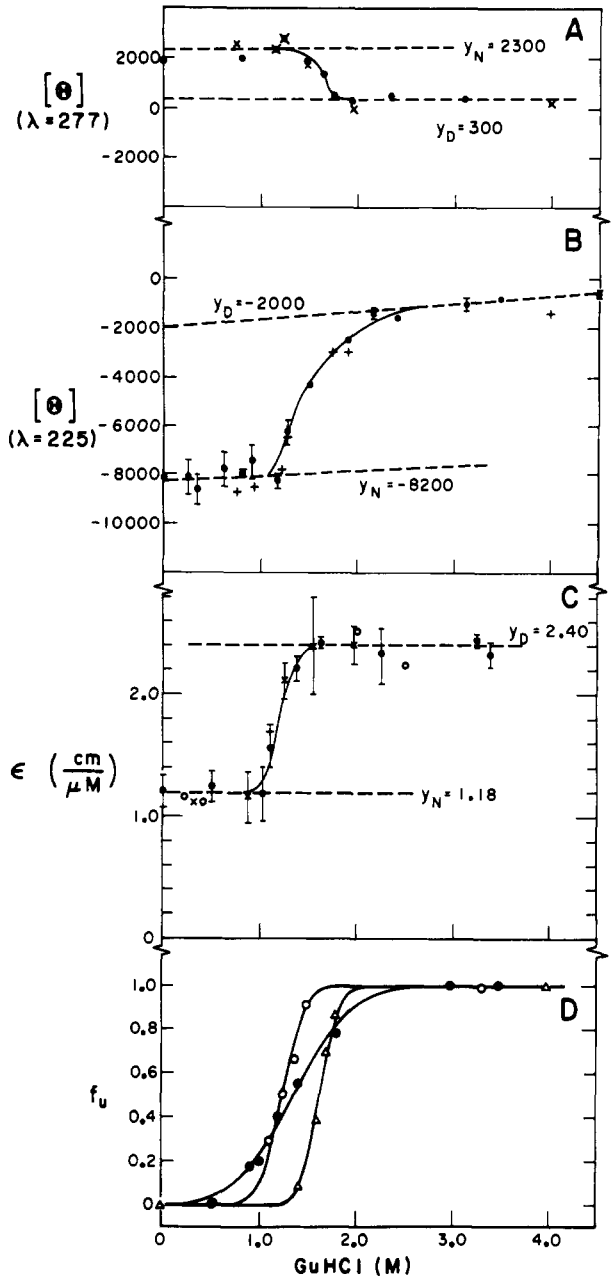


FIGURE 1: Near-UV CD, far-UV CD, and fluorescence measurements as a function of denaturant concentration. y_N and y_D are the intercept values used for the calculation of the fraction of unfolded protein. In panel A, $[\theta]$ is the molar ellipticity per tyrosine from the near-UV CD data collected at λ = 277 nm shown as a function of GuHCl concentration. (×) represents renaturation, and (●) represents denaturation experiments. Each point represents the average of duplicate runs made on the same sample. Panel B contains the molar ellipticity per residue at λ = 225 nm versus [GuHCl]. (●) represents denaturation experiments, and the error bars are derived from standard deviations on multiple samples. (+) indicates renaturation experiments where only duplicate runs of the same sample were done. Panel C depicts fluorescence intensity normalized for protein concentration (micromolar). (●) and error bars are as described above. (×) represents renaturation experiments. (○) and (+) represent denaturation and renaturation experiments where duplicate runs of the same sample were averaged. In panel D, the curves of f_u (eq 1) derived from the data in panels A–C are shown.

GuHCl. The corresponding ΔG is −5.0 kcal/mol. There are four tyrosine residues, three localized around the active site and the fourth at the C-terminus of the protein. The distribution of the fluorescence signal among these four residues is unknown.

The CD spectrum of the native protein in the far-UV shows a bilobed profile with minima at 208 and 225 nm and a local maximum at 217 nm. This is quite different from the case of *E. coli* thioredoxin which has only one lobe with a minimum at 219 nm. The transition midpoint for T4trx is at 1.45 M GuHCl. However, the slope is substantially smaller than in fluorescence. The transition width is about 2 M in GuHCl with a ΔG of −3.0 kcal/mol.

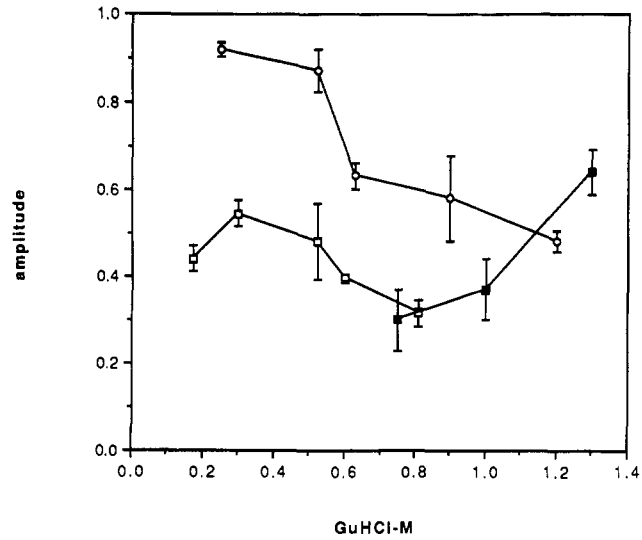


FIGURE 2: Phase amplitudes, α_i, versus final denaturant concentration. The amplitude of the slow phase, α_1, was determined via fluorescent manual mixing studies and the fast phase, α_2, by stopped-flow methods. Protein concentration varied from 2 to 15 μM. (○) α_1 at 25 °C; (□) α_1 at 4 °C; (■) α_2 at 20 °C. Error bars represent standard deviation. The points are simply connected by lines for ease in viewing. Present models do not provide a predicted shape for these curves.

In the CD spectrum in the near-UV, there is a single maximum at 277 nm which corresponds to the tyrosine residues. The midpoint at 1.60 M GuHCl is higher than that of either the far-UV CD or the fluorescence curves. The slope of the curve is slightly sharper than the corresponding fluorescence curve with the transition width again unusually narrow and a corresponding ΔG of −6.5 kcal/mol.

Kinetic experiments were carried out using fluorescence as the monitor of the folding/unfolding transition. The slower refolding phase was studied in manual mixing experiments, but both the fast refolding phase and unfolding required stopped-flow procedures. The rates of the various phases were sufficiently different that semilog plots of the fluorescence traces were essentially linear over the time period of a given experiment. Each run yielded an amplitude and a time constant as described under Materials and Methods.

The amplitudes of the refolding phases are shown in Figure 2. At both 4 and 25 °C, the values of α_1 increase with decreasing final concentration of GuHCl. This is the opposite behavior from that seen in *E. coli* thioredoxin. Although showing some dependence at 4 °C, the time constant τ_1 at 25 °C is essentially independent of the final GuHCl concentration and has a value of about 2.5 min (Figure 3). This general time frame and the lack of strong dependence on denaturant concentration would suggest that τ_1 represents a cis–trans proline isomerization. Previous work has shown that the time range for such a reaction can vary from 0.2 to 7.7 min (Brandts et al., 1975) and is specifically 10–100 s for RNase A (Schmid & Baldwin, 1978). The final GuHCl concentration was varied for both unfolding and refolding.

In the case of unfolding, only one kinetic phase was found. Refolding yielded two phases and $\alpha_1 + \alpha_2 = 1$; thus, apparently no faster phases were missed. The data in Figure 3 do not show the inverted triangle appearance expected for conformational changes [see review by Matthews and Hurle (1987)].

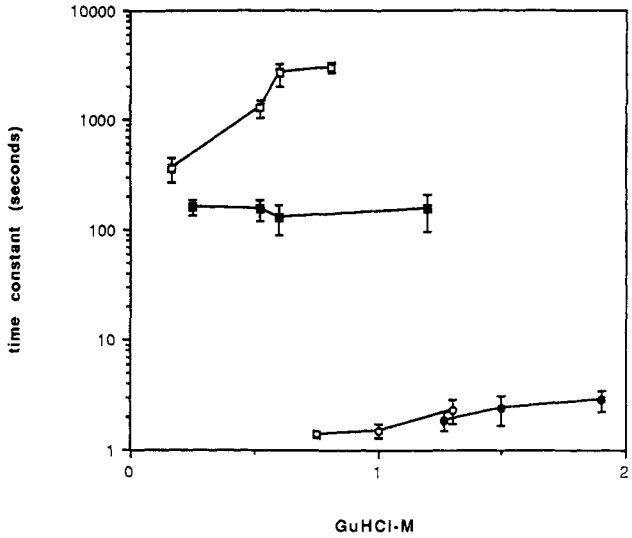


FIGURE 3: Dependence of time constants on final denaturant concentration. The slow phase time constant for refolding, τ_1, at 25 °C (■) and at 4 °C (□); the fast refolding time constant, τ_2, at 20 °C (○); the time constant for unfolding at 20 °C (●). Error bars represent calculated standard deviations.

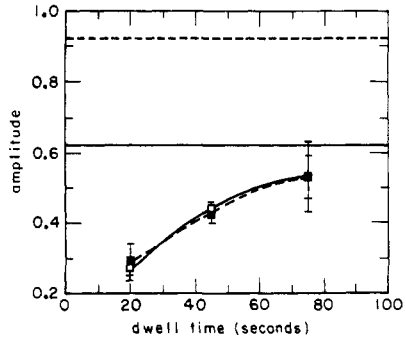


FIGURE 4: Dependence of the phase amplitude, α_1, on dwell time, the time the protein was in concentrated denaturant before it was allowed to refold. The final denaturant concentration was 0.3 M GuHCl for the closed squares (■) and 0.98 M GuHCl for the open squares (□). The horizontal lines indicate the values after 12 h in the final dilute denaturant. Dashed lines correspond to the 0.3 M GuHCl data, and solid lines correspond to the 0.9 M GuHCl. All measurements were done at 25 °C.

However, this result must be characteristic of a local change with regard to the environment around the tyrosines whereas the size-exclusion data discussed below reflect a global change in conformation (see Figure 7).

To further examine the probable cis–trans isomerization, the time constant for the change in amplitude of the slow phase was determined in multimixing experiments. Such experiments were originally reported by Brandts et al. (1975) and used extensively in studies on RNase A (Nall et al., 1978; Schmid & Baldwin, 1978). Samples were unfolded for timed short periods before refolding was initiated. The increase in the amplitude with dwell time is assumed to represent the relaxation of the cis isomer of the native protein to the trans form in the unfolded state. The corresponding reverse reaction for those molecules with a trans proline is then required during the refolding period.

Two such experiments were conducted which differed only in the final denaturant concentration of the refolding step. The two end points were 0.30 and 1.0 M GuHCl. By chance, the data collected in these two runs nearly superimpose, but the asymptotes at infinite dwell time for these two GuHCl concentrations are quite different, and the data actually represent two quite different rates (Figure 4). At 0.3 M GuHCl, the

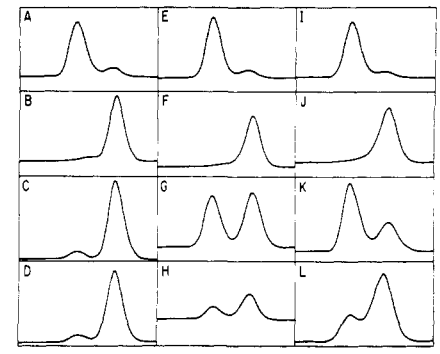


FIGURE 5: Size-exclusion chromatographic profiles showing four sets of experiments. In panels A–D, the column was equilibrated in 1.28 M GuHCl; in panels E–H, 1.56 M GuHCl; and in panels I–L, 1.77 M GuHCl. The first row shows refolding experiments, the second row unfolding experiments, the third row equilibrium experiments, and the fourth row the 30-s denature experiments. All measurements were completed at 4 °C. The ordinate in each case is the optical absorbance of the eluate at 225 nm, and the abscissa shows the elution time from 6 to 10 min.

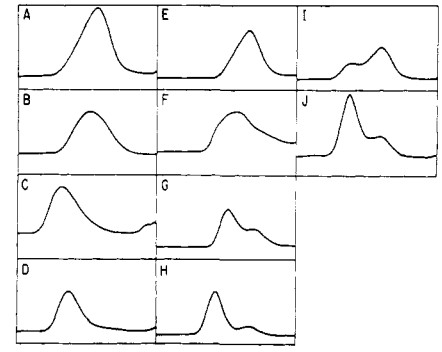


FIGURE 6: Chromatographic elution profiles in the same format as Figure 5. The first vertical set of panels shows profiles near the unfolding transition starting with native proteins in 0 M GuHCl and with final GuHCl concentrations of 1.88, 2.00, 2.10, and 2.3 M for panels A, B, C, and D, respectively. The middle set shows refolding profiles with different final GuHCl concentrations. The protein was denatured for longer than 20 min in 6.2 M GuHCl before injection into the column. The final GuHCl concentrations of the diluted mixtures were 0.20, 0.30, 0.52, and 1.10 M for panels E, F, G, and H, respectively. The pair of panels on the right show profiles for refolding in 1.7 M GuHCl. For the sample in panel I, the protein was denatured for 30 s in 6.2 M GuCHl and in panel J for 900 s in the same solvent.

apparent relaxation rate is 88 s while it is 36 s at 1.0 M GuHCl. The different time constants obtained through these measurements suggest that the cis–trans proline isomerization is linked to a GuHCl-dependent reaction as well, i.e., $U_t \rightarrow I_t$. One would expect that these rates should be identical since they reflect occurrences in the denatured state. We currently do not understand these data. These results do suggest that there must be some other intermediates. One possible problem is that in the experiments where the samples were refolded into 1.0 M GuHCl, those samples may not have been completely out of the transition zone. Thus, these rates may be deceptive if they are coupled to a proline isomerization [see review by Utiyama and Baldwin (1986)].

The elution profiles obtained in the size-exclusion chromatographic runs are shown in Figures 5 and 6. The differing Stokes radius of the native and denatured molecules causes separation of the two forms on the gel columns. In the present study, the column system used gave a peak for the native protein at 8.6 min in the elution cycle while the expanded fully denatured material eluted at 7.4 min. Elution times are somewhat dependent upon denaturant concentration. Within the transition zone, the profile has a definite bimodal shape

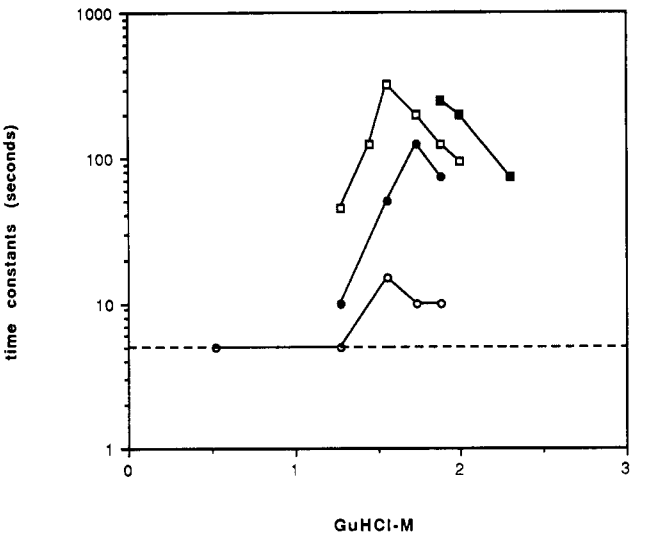


FIGURE 7: Dependence of time constants on final denaturant concentration. Elution profiles were simulated to obtain these time constants according to the method of Shalongo et al. (1987). Open squares (□) indicate the 30-s denature set; closed squares (■) the unfolding set; closed circles (●), the equilibrium set; open circles (○), the refolding set. The apexes for the inverted triangles are 1.60 M GuHCl, 1.73 M GuHCl, and 1.60 M GuHCl, respectively. For the unfolding experiment, the left limb could not be determined. The dashed line marks the time below which reactions are too fast for time constants to be measured reliably.

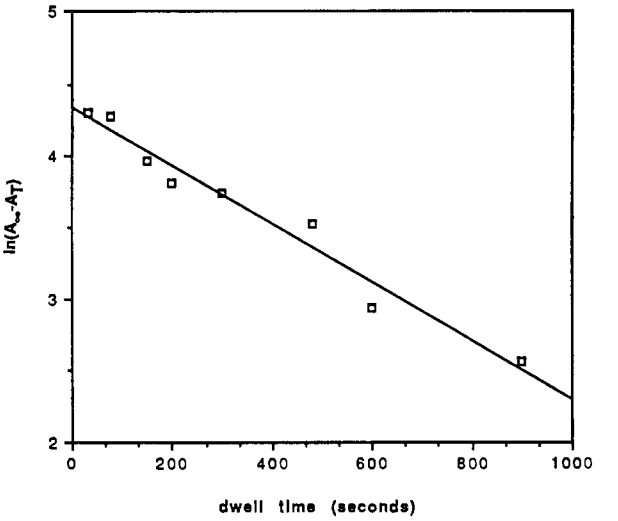


FIGURE 8: Determination of the time constant for the cis–trans proline isomerization. Native protein was placed in 6.2 M GuHCl for varying lengths of time, denoted as dwell time on the abscissa. The sample was then injected into and separated on a column equilibrated in 1.7 M GuHCl. The logarithm of the area under the denatured peak for the given dwell time minus the area at infinite dwell time is shown on the ordinate versus the dwell time. The slope of the line yields a time constant of 560 s.

with a deep valley which suggests that there are two components which are in slow exchange. In the equilibrium profiles (Figure 5, panels C, G, and K), the midpoint of the transition appears to be at about 1.6 M GuHCl. This apparent midpoint may result from contributions of two different refolding pathways which have midpoints of 1.8 and 0.45 M GuHCl (see below). The bimodal profile suggests that there are no significantly populated states which have intermediate volumes between the native and denatured forms. In the profiles of the unfolding reaction (Figure 5, panels B, F, and J), there is no evidence of a denatured peak; i.e., all the protein is in the native form at low denaturant concentrations. In following the unfolding through the transition zone (Figure 6, panels A–D), there is only a single broad peak moving from the native to the denatured position, implying rapid equilibrium on the time scale of the separation between the folded and unfolded states. These data describe the $N \rightarrow U_c$ transition (Scheme II). Simulations of these chromatographs using this three-state model yielded a midpoint of 2.0 M GuHCl. The time constants calculated from this simulation at the various denaturant concentrations are shown in Figure 7.

Refolding experiments suggest that there are at least two paths to arrive at the native structure. One is predominant in higher denaturant concentrations and the other in lower. The two midpoints for this are 1.8 and 0.50 M GuHCl. In higher denaturant concentrations, these profiles are dominated by peaks with elution times of the denatured form, and as the denaturant concentration is lowered, the native form gradually becomes more populated (Figure 5A,E,I). However, there is a burst in formation of native protein around 0.5 M GuHCl (Figure 6E–H). The two midpoints suggest that there are two different pathways involved. The first pathway probably represents $U_c \rightarrow N_c$ whereas the second pathway represents $U_t \rightarrow I_t \rightarrow N_c$. This suggests that I_t does not become stable until the denaturant concentration is low. Originally, Schmid and Baldwin (1976) showed that RNase A has a compact intermediate with a nonnative proline isomer. This mechanism has also been proposed for *E. coli* thioredoxin (Shalongo et al., 1987). Only one profile could be used to calculate τ for the $U_t \rightarrow I_t$ pathway because both species need to be present in order to use the simulation. This time constant is 820 s. The τ's for the $U_c \rightarrow N_c$ transition are shown in Figure 7.

In the multimixing protocol, refolding experiments were performed using protein which had been denatured for varying periods of time. As the dwell time increases, more of the $U_c \rightarrow U_t$ transition occurs so that the area under the denatured peak grows. The time constant τ_{ct} = 560 s was estimated from the slope of comparable data from a series of chromatographic runs (Figure 8). Thirty-second denaturation experiments are simply a subset of the multimixing experiments in which the protein is denatured for such a brief time that essentially only U_c should be populated. A three-state mechanism ($U_t \rightarrow U_c \rightarrow N$) is used to calculate the time constant for the $U_c \rightarrow N$ transition where the time constant for the cis–trans isomerization for $U_t \rightarrow U_c$ is obtained from above. These experiments were done at 1.7 M GuHCl so that the contribution of the $U_t \rightarrow I_t \rightarrow N$ pathway would be negligible. Time constants for that reaction are shown in Figure 7.

Discussion

Both fluorescence mixing studies and size-exclusion chromatography suggest that the slow folding phase is probably a cis–trans proline isomerization. Evidence for this conclusion is the following: (a) the time constant for the slow phase is nearly independent of final denaturant concentration (Figure 3); (b) it is on the minute time scale which is what one would expect for such a reaction; (c) the slow phase is dependent on the dwell time. As the dwell time in the denaturant is increased, the proline will tend to go to the trans form so that both $U_t \rightarrow I_t \rightarrow N$ and $U_t \rightarrow U_c \rightarrow N$ pathways are populated. At equilibrium, the cis:trans ratio in model peptides containing only proline is about 20:80 (Brandts et al., 1975). Therefore, as the dwell time increases, so does the amplitude of the slow phase.

The time constant for the slow phase monitored in fluorescence in T4trx at 25 °C ($U_t \rightarrow I_t \rightarrow N$) is approximately 2.5 min, which is significantly faster than in *E. coli* thioredoxin ($U_t \rightarrow U_c$) which has a τ_1 of approximately 5.0 min. The proline involved in the case of *E. coli* thioredoxin

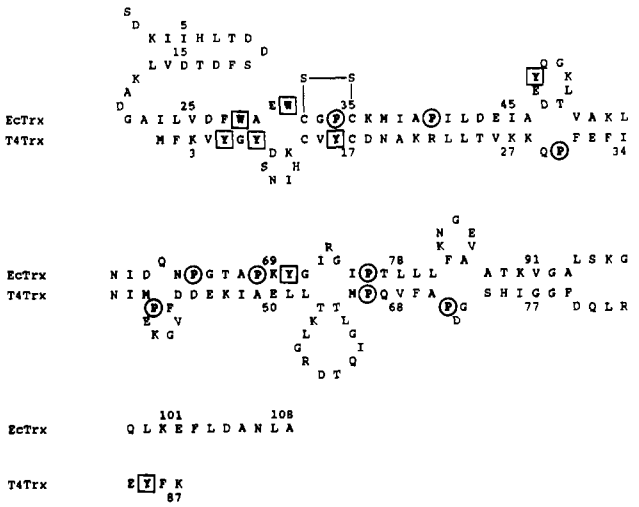


FIGURE 9: Sequences of *E. coli* thioredoxin and T4 thioredoxin are aligned as they are seen in the superimposed 3D structures of the α-carbon chains. In the diagram, the two sequences are closely apposed where there is a good match in the 3D comparison. The structures diverge from each other in the looped regions. These latter parts of the diagram include the actual loops in the secondary structural sense but also parts of helices and β-strands which do not match. The cystine residues, 35 and 37 and 14 and 17, are closely matched, but the sequences of the disulfide loops are different. All of the proline residues are circled. The tryptophans and tyrosines are enclosed in boxes. The two cis prolines-76 and -66 are also in the same position in the 3D structures. There is very little sequence similarity and only four identities in addition to the three just mentioned.

is Pro-76 as shown by the loss of the slow phase when the Pro-76 is changed into an Ala by site-directed mutagenesis (Kelley & Richards, 1987). Pro-76 is directly in front of the fourth β-strand and is adjacent to a β-branched residue, Ile. The analogous proline in T4 thioredoxin is Pro-66, which is also cis and in front of the fourth β-strand (Figure 9). However, Pro-66 is adjacent to Met-65 and Gln-67, neither of which are β-branched (Söderberg et al., 1978). Therefore, the slower τ_1 in *E. coli* thioredoxin might be due to difficulty in isomerizing due to a steric effect caused by the proximity of a β-branched group.

As can be seen in Figure 9, with the exception of one pair of prolines (76 and 66), the reporter groups affecting the measurements in this study are not located in identical places in the 3D structures of the *E. coli* and T4 thioredoxins.

The slow folding phase in T4Trx is most likely due to the $I_t \rightarrow N$ proline isomerization (Scheme II). This is different from the *E. coli* case, in which the slow phase is due to the $U_c \rightarrow U_t$ isomerization. The critical point to focus on is the basis for the fluorescence changes in T4 and *E. coli* thioredoxin. Fluorescence increases upon denaturation in *E. coli* thioredoxin due to the removal of the quenching from the disulfide bond. The two tryptophans in EcTrx are 7–8 Å away from the center of the disulfide bond. However, disulfide quenching does not play an important role in T4Trx fluorescence. Two of the four tyrosines are also within 8 Å (S. Katti, personal communication), but there is little change in fluorescence intensity in the native reduced versus native oxidized forms (K. L. B. Borden, unpublished results). Thus, the fluorescence increase upon denaturation is probably due to exposure of tyrosyl chromophores rather than removal of quenching groups. In such a case, one would expect the fluorescence of the tyrosyl groups to be perturbed by the proline isomerization in the more compact form where the critical proline is actually nearby. Intuitively, this seems more likely because in the unfolded state one expects the tyrosine fluorescence to be unaffected by the $U_c \rightarrow U_t$ reaction, assuming, of course, that the denatured state is actually an expanded random coil. However, in the case of RNase A, Schmid (1981) did observe the $U_c \rightarrow U_t$ reaction via tyrosine fluorescence.

From the size-exclusion data, it is known that the midpoint in the $U_t \rightarrow I_t$ transition occurs at 0.50 M GuHCl (Figure 6E–H). From manual mixing experiments, the midpoint of α_1 at 4 °C is 0.5 M GuHCl and, as stated previously, increases at lower denaturant concentrations. This is consistent with the model because, as a native-like structure, the stability of I_t should increase in lower denaturant concentrations (Figure 2). The final slow phase amplitude, obtained from fluorescence data, is significantly higher at 25 °C than at 4 °C. This fact suggests that I_t is more stable at 25 °C than at 4 °C; i.e., I_t is cold-denatured in this temperature range (Privalov et al., 1986).

Both the fluorescence kinetic and size-exclusion chromatography data are adequately described by a four-state box mechanism (Scheme II). A cis–trans proline isomerization is important to folding, and there are two pathways from which the unfolded protein can arrive at the final native form. This mechanism is a minimum model. The complexities demonstrated by the equilibrium curves suggest that there are probably more intermediates; however, there is no kinetic evidence at this time to indicate where in the pathway such intermediates may exist. Furthermore, a similar study on T4 thioredoxin at pH 7.0 and 4 °C indicates that another mechanism besides the one shown in Scheme II may be involved simultaneously (E. Stellwagen, personal communication).

The equilibrium data (Figure 2) indicate that the mechanism is probably more complex than our Scheme II. The different midpoints for fluorescence and near-UV CD data indicate that there are at least two populations of tyrosines. Both curves are unusually sharp and are completed within the transition width of the far-UV CD curve. The latter data almost certainly represent changes in the secondary structure of the main chain. Thus, it would appear that groups of tyrosyl chromophores are being exposed to solvent at different points during the melting of the secondary structures. This implies the existence of separate domains in the T4 structure which unfold at distinctly different denaturant concentrations. The main chain changes accompanying these two events are seen in the continuous broad curve for the far-UV CD. One of these subsets of tyrosines is affected by the $I_t \rightarrow N_c$ transition as seen in the manual mixing fluorescence experiments.

A number of proteins have been shown to have noncoincident equilibrium curves when monitored by more than one spectroscopic probe, for example, carbonic anhydrase (Wang & Tanford, 1973), β-lactoglobulin (Ananthanaraynan & Ahmad, 1977), α-lactalbumin, (Kawajima et al., 1976; Dolgikh et al., 1985), and bovine growth hormones (Brems et al., 1985; Havel et al., 1986; Brems & Havel, 1989). Since the different physical measurements reflect different structural parameters, noncoincident curves require one or more intermediate states. This may be most clearly illustrated by probes that monitor global changes in structure such as far-UV CD and size-exclusion chromatography and probes that monitor local changes as seen by fluorescence and near-UV CD.

ACKNOWLEDGMENTS

Many thanks to Dr. Earle Stellwagen's group for allowing us to use their size-exclusion chromatography setup and also for many helpful discussions regarding the manuscript. Special thanks to M. V. Jagannadham and W. Shalongo for their technical assistance during the stay in Iowa. Many thanks

are also due to Dr. Todd Schuster at the University of Connecticut for allowing us to use his DG110 stopped-flow fluorometer and for the technical advice that he and John Philo provided. Our thanks to G. Johnson for advice and help with the limping DG110 stopped-flow unit at Yale. Thanks to D. LeMaster for helpful discussions. Our special thanks to Hans Eklund for access to the X-ray coordinates of T4 thioredoxin at the current state of refinement. We acknowledge A. Johnson, T. Mouning, and J. Mouning for their help with manuscript preparation and S. Katti for help with computer graphics.

References

Ananthanaraynan, V. S., & Ahmad, F. (1977) *Can. J. Biochem. 55*, 239–243.

Berglund, O., & Sjöberg, B.-M. (1970) *J. Biol. Chem. 245*, 6030–6035.

Brandts, J. F., Halvorson, H. R., & Brennan, M. (1975) *Biochemistry 14*, 4953–4963.

Brems, D. N., & Havel, H. A. (1989) *Proteins: Struct., Funct., Genet. 5*, 93–95.

Brems, D. N., Plaisted, S. M., Havel, H. A., Kauffman, E. W., Stodola, J. D., Eaton, L. C., & White, R. D. (1985) *Biochemistry 24*, 7662–7668.

Dolgikh, D. A., Abaturov, L. V., Bolotina, I. A., Brazhnikov, E. V., Bychkova, V. E., Gilmanshin, R. I., Lebedev, Yo. O., Semisotnov, G. V., Tiktopulo, E. I., & Ptitsyn, O. B. (1985) *Eur. Biophys. J. 13*, 109–121.

Endo, S., Saito, Y., & Wada, A. (1983) *Anal. Biochem. 131*, 108–120.

Havel, H. A., Kaufman, E. W., Plaisted, S. M., & Brems, D. N. (1986) *Biochemistry 25*, 6533–6538.

Holmgren, A., Söderberg, B.-O., Eklund, H., & Branden, C.-I. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 2305–2309.

Katti, S. K., LeMaster, D. M., & Eklund, H. (1989) *J. Mol. Biol.* (in press).

Kelley, R. F., & Stellwagen, E. (1984) *Biochemistry 23*, 5095–5103.

Kelley, R. F., & Richards, F. M. (1987) *Biochemistry 26*, 6765–6774.

Kelley, R. F., Wilson, J., Bryant, C., & Stellwagen, E. (1986) *Biochemistry 25*, 728–732.

Kuwajima, K., Nitta, K., Yoneyama, M., & Sugai, S. (1976) *J. Mol. Biol. 106*, 359–373.

LeMaster, D. M. (1986) *J. Virol. 59*, 759–760.

LeMaster, D. M., & Richards, F. M. (1988) *Biochemistry 27*, 142–150.

Lin, L.-N., & Brandts, J. F. (1983) *Biochemistry 22*, 553–559.

Mathews, C. R., & Hurle, M. R. (1987) *BioEssays 6*, 254.

Nall, B. T., Gavel, J.-R., & Baldwin, R. L. (1978) *J. Mol. Biol. 118*, 317–330.

Privalov, P. L., Griko, Yu. V., Venyaminov, S. Yu., & Kutyshenko, V. P. (1986) *J. Mol. Biol. 190*, 487–498.

Schmid, F. X. (1981) *Eur. J. Biochem. 114*, 105–109.

Schmid, F. X., & Baldwin, R. L. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 4764–4768.

Shalongo, W., Ledger, R., Jagannadham, M. V., & Stellwagen, E. (1987) *Biochemistry 26*, 3135–3141.

Söderberg, B.-O., Sjöberg, B.-M., Sonnerstam, U., & Branden, C.-I. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 5827–5830.

Utiyama, H., & Baldwin, R. L. (1986) *Methods Enzymol. 131*, 51–70.

Wong, K. P., & Tanford, Ch. (1973) *J. Biol. Chem. 248*, 8518–8523.

Pulsed Electron Paramagnetic Resonance Studies of the Interaction of Mg-ATP and D_2O with the Iron Protein of Nitrogenase†

T. V. Morgan,‡ J. McCracken,§ W. H. Orme-Johnson,‖ W. B. Mims,⊥ L. E. Mortenson,*,‡ and J. Peisach§

University of Georgia, Center for Metalloenzyme Studies, Athens, Georgia 30602, Department of Molecular Pharmacology, Albert Einstein College of Medicine, Bronx, New York 10461, Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, and Consultant, Exxon Research and Engineering Company, Annandale, New Jersey 08801

Received October 26, 1989

ABSTRACT: Mg-ATP binds to the iron protein component of nitrogenase. The magnetic field dependence of the linear electric field effect (LEFE) in pulsed EPR is consistent with a single 4Fe-4S cluster. The LEFE is virtually unaltered when Mg-ATP is bound. Electron spin echo envelope modulation techniques were employed to evaluate the possibility of a magnetic interaction between ^{31}P of Mg-ATP and the Fe-S center of the iron protein. None was detected. However, weak modulations possibly attributable to peptide ^{14}N were seen, and these were slightly shifted by Mg-ATP addition. Further, protons in the vicinity of the Fe-S cluster of the protein readily exchange with D_2O, and this process is unaffected by Mg-ATP.

Nitrogenase catalyzes the reduction of dinitrogen to ammonia (Burgess, 1984; Orme-Johnson, 1985). It can be separated into two components, one an iron–sulfur protein here termed the iron protein, consisting of two equivalent 30-kDa subunits, and a larger protein containing an iron–molybdenum cofactor (Nakos & Mortenson, 1971). The current view is that the iron–sulfur protein contains a 4Fe-4S cluster bound at an interface between both subunits (Anderson & Howard, 1984). Enzyme catalysis depends on the transfer of electrons by the reduced iron–sulfur protein to the iron–molybdenum protein, concurrent with hydrolysis of bound Mg-ATP (Thorneley et al., 1979).

† This work was supported by U.S. Public Health Service Grants GM 40067-02 to L.E.M. and GM 40168 and RR 02583 to J.P.

‡ University of Georgia.
§ Albert Einstein College of Medicine.
‖ Massachusetts Institute of Technology.
⊥ Exxon Research and Engineering Company.

0006-2960/90/0429-3077$02.50/0