Proline Isomerization-Independent Accumulation of an Early Intermediate and Heterogeneity of the Folding Pathways of a Mixed α/β Protein, *Escherichia coli* Thioredoxin†

Roxana E. Georgescu,‡,§ Jian-Hua Li,‡ Michel E. Goldberg,§ Maria Luisa Tasayco,‡ and Alain F. Chaffotte*,§

Biochemistry Division, Chemistry Department, City College of the City University of New York, New York, New York 10031, and Unité de Biologie Cellulaire, Institut Pasteur, Paris 75015, France

Received March 4, 1998; Revised Manuscript Received May 18, 1998

ABSTRACT: Oxidized *Escherichia coli* thioredoxin (Trx) is a small protein of 108 residues with one disulfide bond (C32−C35 essentially involved in the activity) and no prosthetic moieties, which folds into a structural motif containing a central twisted β-sheet flanked by helices that is found in many larger proteins. The kinetics of refolding of Trx in vitro have been investigated using a newly developed active site titration assay and continuous or stopped-flow (SF) methods in conjunction with circular dichroism (CD) and fluorescence (Fl) spectroscopy. These studies revealed the presence of early folding intermediates with "molten globule or pre-molten globule" characteristics. Measurements of the ellipticity at 222 nm indicated that about 68% of the total change associated with refolding occurred during the dead time (4 ms) of the stopped-flow instrument, suggesting the formation of substantial secondary structure. The reconstruction of the far-UV CD spectrum of the burst intermediate using combined continuous and stopped-flow methods showed the formation of a defined secondary structure that contains more β-structure than the native state. Kinetic measurements using SF far-UV CD and Fl over a wide range (0.087−6 M) of GuHCl concentrations at two temperatures (6 and 20 °C) demonstrated that the population formed during the 4 ms dead time contained multiple species that are stabilized mainly by hydrophobic interactions and undergo further folding along alternative pathways. One of these species leads directly and rapidly to the native state as demonstrated by active site titration, while the two others fold into a fourth intermediate that is slowly converted to the native protein. Double-jump experiments suggest that the heterogeneity in folding behavior results from proline isomerizations occurring in the unfolded state. Conversely, the accumulation of the burst intermediate does not depend on proline isomerizations.

Understanding the process of protein folding is one of the major challenges of modern structural and molecular biology. Elucidation of the structure of intermediates formed during the folding process has been the goal of many experiments, but the rapidity and high cooperativity of this process make it difficult to study the structural and thermodynamic properties of folding intermediates. This is particularly true at early stages (a few milliseconds or less) of the folding process, where it has been proposed that the structures formed are not persistent and that the corresponding conformers are highly dynamic and fluctuate rapidly (*1*). Yet, time-resolved fluorescence and CD[1] kinetic experiments have revealed the existence of early folding transients for many small proteins such as cytochrome *c* (*2*), staphylococcal nuclease (*3*), ubiquitin (*4*), apomyoglobin (*5*), lysozyme (*6*−*8*), ribonuclease A (*9*), etc., suggesting that folding develops through a limited number of pathways with defined intermediates (*10*).

Because of its well-defined mixed α/β secondary structure elements, stability, small size, and prokaryotic origin, which facilitates isotopic labeling and expression of mutant proteins, *Escherichia coli* thioredoxin provides an excellent model system for the study of protein folding and dynamics. Its three-dimensional structure was determined by X-ray diffraction (*11*) and multidimensional NMR spectroscopy (*12*). These studies show that the 108-residue polypeptide chain folds into a compact globular conformation with a central twisted, five-stranded β-sheet packed between the three α-helices observed in solution by NMR. Its folding mechanism has already been in part determined with kinetic studies (*13*, *14*). Using stopped-flow fluorescence measurements and manual mixing with fluorescence and CD, these authors identified three phases during the refolding of oxidized thioredoxin in 2 M GuHCl. The slowest of these phases was attributed to the trans−cis isomerization of Pro 76 (*15*).

In the work reported here, we focused our efforts on characterizing the early events occurring during thioredoxin

† This work was supported by the RCMI grant from NIH to CCNY, Grants MCB-9507255 and INT-9600006 from NSF to M.L.T., who is an NSF Career Awardee, and funds from the Institut Pasteur, the Centre National de la Recherche Scientifique (URA 1129), and the Université Paris 7.

* To whom correspondence should be addressed.

‡ City College of the City University of New York.

§ Institut Pasteur.

[1] Abbreviations: Trx, oxidized *E. coli* thioredoxin; SF, stopped-flow; CD, circular dichroism; GuHCl, guanidinium hydrochloride; DTNB, 5,5′-dithiobis(2-nitrobenzoic acid); ANS, 8-anilino-1-naphthalenesulfonate.

S0006-2960(98)00508-X CCC: $15.00
Published on Web 06/20/1998

folding, using direct kinetic measurements based on rapid mixing methods. The results we report indeed show that as early as 4 ms a burst intermediate is already formed. Its far-UV CD spectrum reflects the presence of large amounts of secondary structure, but the amounts of α-helix and β-structure it contains differ considerably from those present in native thioredoxin. The effects of GuHCl and temperature on the CD spectrum of the burst intermediate(s) will be reported and will be used to gain some insight into the origin of its stability. Double-jump experiments aimed at investigating the possible role of proline isomerization during the fast folding phases will be described. The results will be discussed in terms of a multistep folding process that involves parallel pathways leading to the native state.

MATERIALS AND METHODS

Materials and Protein Purification

Ultrapure guanidine hydrochloride (GuHCl) was purchased from ICN Biomedicals Inc. and used without further purification. Dithiothreitol, bovine serum albumin, EDTA, and NADPH were from Sigma Chemical Co. (St. Louis, MO), and 5,5′-dithiobis(2-nitrobenzoic acid) (DTNB) was from Aldrich. *E. coli* thioredoxin reductase was purchased from IMCO Corp. Ltd. All other chemicals were reagent grade. The buffer used in all experiments was 10 mM potassium phosphate at pH 7.0 (KP_i buffer). Wild type Trx was overexpressed in *E. coli* JF521 (gift from J. Fuchs and C. Woodward) and purified using previously described gel filtration and ion exchange chromatographic procedures (*16*). Protein purity was verified by the presence of a single band on Coomassie blue-stained SDS−polyacrylamide gels (*17*), by amino acid analysis, and by electrospray mass spectrometry. GuHCl concentrations were determined before and after refolding experiments using the refractive index measurement (*18*), performed on an Abbe refractometer. Protein concentrations were determined spectrophotometrically, using an extinction coefficient of 1.37×10^4 M^{-1} cm^{-1} (*19*).

CD Spectroscopy

Equilibrium and time-resolved CD experiments were performed on a CD6-circular dichroism spectropolarimeter (Jobin-Yvon ISA, Longjumeau, France) equipped with a SFM-3 stopped-flow module (Bio/Logic, Claix, France). The stopped-flow unit and the observation cell (path length of 1.5 mm) were thermostated at the working temperature appropriate for each experiment by circulating water from a temperature-controlled water bath.

(a) For investigations on the stability of the burst intermediate, Trx denatured in 4 M GuHCl was renatured at two different temperatures (6 and 20 °C) by rapid mixing (60 ms injection time) of 30 μL of protein with 570 μL of KP_i buffer containing GuHCl at concentrations that covered the range between 0.2 and 4 M. The dead time of the instrument was 3 ms. The ellipticity at 222 nm was recorded, using a filtering time constant less than or equal to the sampling interval. The total acquisition time varied from 2 to 100 s depending on the final GuHCl concentration. For each time base, 50 transients of 4001 data points each were averaged. Controls with refolding buffer and GuHCl solutions were recorded for correcting the kinetic traces of any instrument offsets.

(b) For the reconstruction of the burst intermediate spectrum, at 20 °C, we used a combined continuous and stopped-flow mixing procedure comprising three phases. First, the plateau of the previous kinetic trace was recorded during a 600 ms noninjective phase. During a subsequent 400 ms continuous-flow phase, 0.2 mL of unfolded Trx was mixed with 3.8 mL of refolding buffer. The flow was then stopped, and data were recorded over the course of 4 s. During these three phases, data points were collected at 5 ms intervals, using a 5 ms filtering constant. The ellipticity during the continuous-flow phase was determined by averaging 80 data points and was compared with the ellipticity calculated from the back-extrapolation of the kinetic traces recorded during the interrupted flow phase. The flow rate during the injection phase was 10 mL/s, resulting in an effective folding time of ∼2.7 ms for the protein observed during the injection phase (delay between the mixer and the center of the observation cell). Using this procedure, nine successive combined traces were averaged at each wavelength (step of 2 nm). At each wavelength, the CD signal for the refolding buffer alone was used as a baseline and was subtracted from all other data to remove instrument offsets.

(c) Equilibrium CD spectra in the far-UV CD (204−240 nm) region of both unfolded (in 4 M GuHCl) and refolded (in 0.2 M GuHCl) samples of Trx (22 μM) were recorded at 20 °C and pH 7.0 in a 1.5 mm optical path cell, using 1 nm scanning steps, a 2 s integration time, and a 2 nm bandwidth. For each sample, five accumulations were averaged and the spectrum of the corresponding solvent was subtracted.

Stopped-Flow Fluorescence Spectroscopy

Kinetic measurements were conducted at different temperatures using a Bio/Logic spectrofluorimeter equipped with a SF-3 stopped-flow module and a 1.5 mm × 1.5 mm observation cell. A 150 W xenon−mercury lamp and a monochromator were used for excitation at 295 nm (2 mm slit). The fluorescence emission was measured using a high-pass filter with a 325 nm cutoff. Two photomultipliers were used, one for detection of the fluorescence emission and the other as a reference for monitoring the intensity of the excitation signal. The ratio of the two signals was either recorded directly through a Bio/Logic dual amplifier or calculated for each couple of data points by the Bio-Kine software (Bio/Logic) acquisition system in the ratio mode. In some experiments, the output of the detection photomultiplier was sent simultaneously to two different amplifiers and recorded separately, on two distinct channels of the acquisition system, with different time scales and filtering time constants.

(a) Stopped-Flow Fluorescence Stability Experiments. Kinetic measurements of Trx refolding at 6 and 20 °C were obtained as described in CD Spectroscopy by 20-fold dilution (60 ms injection time) of denatured Trx (2 mg/mL in 4 M GuHCl) with KP_i buffer containing different amounts of GuHCl (for achieving final concentrations ranging between 0 and 4 M). The time scales chosen were 0−5 and 0−20 s, 0−10 and 0−50 s, or 0−100 s, depending on the residual GuHCl concentration. In some cases, two different couples of time constants were used in experiments with the same GuHCl concentration, which allowed us to reliably paste

kinetic traces covering a large time window. All the solutions were filtered and extensively degassed immediately before they were used. Controls with GuHCl solutions at the working concentration have also been recorded. For each GuHCl concentration, 10−30 transients of 1000 data points each were averaged and pasted for different acquisition times.

(b) Activation Energy Determination. The refolding of Trx was initiated by rapid 20-fold dilution (injection time of 50 ms) of denatured Trx (2 mg/mL in 4 M GuHCl) with KP_i buffer and the progress recorded at different temperatures between 6 and 30 °C, using appropriate acquisition times; for each temperature, 5−20 transients of 1000 points were averaged and pasted. For lower temperatures, additional traces were recorded in the manual mode using a thermostated PTI spectrofluorimeter. The excitation and emission band-passes were 2 and 5 nm, respectively. The excitation and emission wavelengths were 295 and 350 nm, respectively. The total acquisition time was 1800 s (one point per second). The manual and stopped-flow kinetic traces were then pasted, after normalization.

(c) Binding of ANS. ANS binding was monitored after 20-fold dilution of denatured Trx (2 mg/mL in 4 M GuHCl) with KP_i buffer containing 0.1 mM ANS. Two final residual GuHCl concentrations (0.2 and 0.5 M) and two temperatures (6 and 20 °C) were used. Stopped-flow kinetic traces were acquired by recording the fluorescence (excitation at 390 nm, emission through a high-pass filter with a cutoff at 420 nm) on different time scales as indicated above, to cover the whole refolding process; controls for the native and denatured protein were recorded, at the specific residual GuHCl concentrations.

(d) Double-Jump Experiments. Classical double-jump experiments (native−unfolded−native) were performed by manual mixing (for the slow phase) and stopped-flow measurements. The experimental details will be described in the Results. To check for a possible dimerization of the native protein at the high concentration (10 mg/mL) used as the starting concentration in the double-jump kinetics, a sedimentation equilibrium experiment at 27 000 rpm and 20 °C was performed in a XLA analytical ultracentrifuge (Beckman), using the absorption scanning system at 308 nm. No association of the Trx molecules could be detected (data not shown).

(e) Activity and Kinetics of Reactivation Measurements. For steady-state kinetics, thioredoxin-catalyzed reduction of DTNB was monitored in the presence of *E. coli* thioredoxin reductase and NADPH (*20*).

For kinetic measurements, the amount of native Trx was determined by active site titration, monitoring the decrease in NADPH fluorescence (excitation at 350 nm and emission at 450 nm) during the reaction. Assay mixtures, extensively deaerated before the measurements, contained 0.1 M phosphate buffer (pH 7.0), 2 mM EDTA, 50 μg/mL BSA, and 24 μM NADPH in a final volume of 0.5 mL. Thioredoxin reductase was added at a final concentration of 0.2 μM; the denatured protein in 4 M GuHCl or native thioredoxin was added at a final concentration of 6 μM (it ensured a reasonable reaction time and a good signal decay). The final concentration in GuHCl in all the experiments was 87 mM. A set of titration assays with different final concentrations of native thioredoxin ranging between 0.6 and 6 μM was performed, to calibrate its fluorescence amplitude as a function of the protein concentration. The dead time in all cases was 6−11 s. NADPH and the native or denatured thioredoxin displayed a flat baseline for 1900 s. In the absence of thioredoxin, NADPH was oxidized by thioredoxin reductase at a rate 10 times slower than that in the presence of the native protein.

Manual multimixing kinetic reactivation measurements were performed by titrating, as indicated above, the formation of native thioredoxin active sites at different times during the unfolding−refolding process.

Data Analysis

The kinetic data were resolved into phases as a sum of individual exponential terms using nonlinear least-squares methods for fitting to the equation

$$X_t = \sum_i X_i e^{-k_i t} + X_\infty \quad (1)$$

where X_t is the fluorescence or ellipticity at time t, X_i is the total amplitude of the change associated with phase i, k_i is the apparent rate constant of phase i, and X_∞ is the amplitude at equilibrium. The number of phases was increased until no significant improvement of the fit was observed as judged from the residual errors. The signals of intermediate species were extracted from the general phase amplitudes (X_i) according to the kinetic model proposed in this paper (Scheme 3).

The transition unfolding curves in GuHCl for the burst and the rapid folding intermediates of Trx were analyzed assuming a two-state mechanism for unfolding (*21*). All parameters were determined by nonlinear least-squares fitting of the following equation, using the Fig.P software (Biosoft):

$$S([D]) = S_{n,0} + k_n[D] + \{S_{u,0} - S_{n,0} + (k_u - k_n)[D]\}\frac{K_0 e^{m[D]/RT}}{1 + e^{m[D]/RT}} \quad (2)$$

where $S([D])$ is the value of the specified optical parameter at denaturant concentration [D], $S_{n,0}$ and $S_{u,0}$ represent the calculated values of S extrapolated at zero denaturant for the native and unfolded states, respectively, and k_u and k_n are the slopes of the unfolded and native baselines, respectively. K_0 represents the equilibrium constant of the transition, and m represents the slope of the free energy change versus denaturant concentration, assuming a linear relationship between the free energy of unfolding and denaturant concentration [D] (*18*).

The far-UV CD burst phase spectrum reconstructed from kinetic refolding experiments and the spectra of the native or denatured thioredoxin were analyzed for the content in secondary structure using the following programs: Varselec (*22*), Selcon (*23*), Contin (*24*), including four denatured peptides (*25*), K2D (*26*), and Prosec (*27*).

RESULTS

Unfolding Equilibrium. The GuHCl-induced unfolding transition of oxidized Trx, monitored by tryptophan fluorescence and by ellipticity at 222 nm at 20 °C and pH 7.0, occurs between 2 and 3.5 M GuHCl with a midpoint concentration (C_m) of 2.8 ± 0.1 M. Equilibrium CD

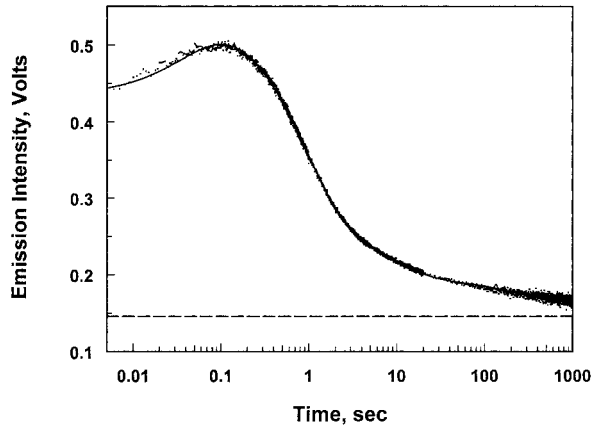


FIGURE 1: Trx refolding kinetics monitored by fluorescence. Time course of refolding of *E. coli* thioredoxin in 87 mM residual GuHCl (40-fold diluted denatured protein in 3.5 M GuHCl) and 20 °C, monitored by fluorescence emission of ANS as the fluorescent probe. The solid line represents the best fit of three averaged collated kinetic traces, using four exponentials. The dashed line represents the control signal for 0.1 mM ANS.

measurements also showed that the ellipticity at 222 nm is virtually independent (it behaves as a denatured baseline) of the GuHCl concentration above 3.4 M. Thus, the protein can be considered fully unfolded at 4 M GuHCl. This concentration was therefore chosen in these studies as the starting concentration for most refolding experiments. Similarly, the final GuHCl concentrations selected for most refolding experiments, 0.087 and 0.2 M, lie far below the unfolding transition.

Refolding Kinetics Monitored by CD and Fluorescence. In a preliminary set of experiments, the kinetic parameters for the complete refolding process of Trx in the presence of 87 mM residual GuHCl were determined from the kinetic traces obtained when recording the far- and near-UV CD, the intrinsic fluorescence (*28*), and the binding of ANS (Figure 1). The results obtained are summarized in Table 1. These data show that, except for the near-UV CD where the folding process could be entirely described by three phases, all the time courses exhibited a burst followed by four phases. These results differ from those reported previously (*13*, *14*) which showed only three relaxation times for the refolding of Trx.

The presence of a significant burst amplitude (67% in far-UV CD and 30% in fluorescence) and of four rather than three observable phases raised some questions about the species formed during the burst phase, its structural features, and finally its place in the folding pathway proposed in previous studies (*14*, *29*, *30*). We therefore attempted to characterize in more detail the species formed during the burst and to carefully reanalyze the kinetics of Trx folding. However, because of its high rate constant, the very rapid kinetic phase, which immediately follows the burst, could not be monitored accurately under the experimental conditions used above. We therefore chose to carry out further studies on the early phases of Trx refolding by performing stopped-flow experiments in the presence of 0.2 M GuHCl, the lowest GuHCl concentration at which the very rapid kinetic phase could be easily detected and characterized.

Far-UV CD Spectrum of the Burst. The far-UV CD spectrum of the burst was reconstructed from individual refolding kinetic traces acquired at different wavelengths (each 2 nm between 212 and 240 nm), using a combined continuous and stopped-flow method (see Materials and Methods). Figure 2A shows a typical complete recording (1000 data points) of such kinetics, which contains 80 data points acquired during the 400 ms continuous-flow phase, followed by the trace in the stopped-flow mode. The refolding reaction showed a burst (48% of the total change at 222 nm) followed by a rapid decrease in the ellipticity which, under the conditions used (0.2 M residual GuHCl), could easily be monitored even during the first milliseconds after arresting the flow. On the time scale of these experiments (5 s), a curve derived from a two-phase kinetic model could be fit very well to the kinetic trace obtained in the stopped flow (Figure 2A). Similar experiments were repeated at different wavelengths, and analysis of the kinetics (as a sum of exponentials for providing the rate constant and amplitude of each phase) was performed. At each wavelength investigated, a satisfactory fitting with two exponentials was obtained. As shown in Figure 2A for the kinetics at 222 nm, the "initial" ellipticity obtained at each wavelength by back extrapolation of the fitted curve was very close to the ellipticity observed for the burst in the continuous flow. From the amplitudes obtained at each wavelength, we could reconstruct the spectrum describing the wavelength dependence of the ellipticity change associated with each phase (data not shown). In the case of the slow phase, the same rate constant [$(2.17 \pm 0.29) \times 10^{-3}$ s^{-1}] was found by monitoring refolding of denatured Trx at different wavelengths, and the spectrum reconstructed from the final values of the kinetic traces closely matched that of the native protein (*28*).

The far-UV CD spectrum of the species formed during the burst phase (Figure 2B, open circles) was first analyzed by assuming that these species consisted of a mixture of only the native and unfolded Trx. That the spectrum of the burst phase could not be fit by a linear combination of the spectra of native and denatured Trx (Figure 2B, dash−dot line) ruled out this hypothesis. We then tried to interpret quantitatively the far-UV CD spectrum of the burst intermediate. This was possible because of the absence of any detectable near-UV ellipticity in the burst intermediate, which indicated that the aromatic and cystinyl side chains are still in a symmetrical environment. Thus, unlike the situation of native Trx where the far-UV CD spectrum is strongly distorted by contributions from these side chains, the spectrum of the burst intermediate reflects only its backbone conformation. This conclusion was used to interpret the burst far-UV CD spectrum in terms of secondary structure. It was first assumed that the burst species might consist of a mixture of unfolded Trx and of "molten globule-like" molecules (*31*) with a native-like secondary structure and no tight packing of its side chains. This assumption was ruled out since the far-UV CD spectrum of the burst intermediate could not fit by a linear combination of the spectrum of the denatured Trx and that predicted for the contribution of the backbone alone in the native protein based on its known secondary structure (Figure 2B, thick continuous line). We then attempted to characterize the secondary structure of the burst intermediate with different secondary structure prediction programs (Prosec, Contin, Selcon, and K2D). As shown in Table 2, all the deconvolution methods gave a secondary structure substantially different from that of the native protein

Table 1: Kinetic Parameters for Refolding of Denatured Oxidized *E. coli* Thioredoxin[a]

	far-UV CD[b]		near-UV CD[c]		fluorescence[d]		
phase	k (s^{-1})	relative amplitude (%)	k (s^{-1})	relative amplitude (%)	k (s^{-1})	relative amplitude (%)	ANS[d] k (s^{-1})
BP	>400	68 ± 2	−	−	>400	30.4 ± 0.9	>400
VRP	35 ± 2	22.8 ± 0.9	33 ± 2	19 ± 1	33 ± 2	6.9 ± 0.2	29 ± 1
RP	0.62 ± 0.03	12.4 ± 0.3	0.73 ± 0.03	45 ± 2	0.429 ± 0.003	39.8 ± 0.2	1.72 ± 0.03
MP	0.05 ± 0.02	1.6 ± 0.4	0.04 ± 0.02	36 ± 1	0.04 ± 0.003	6.3 ± 0.3	0.175 ± 0.031
SP[a]	2.5×10^{-3}	−6 ± 2	−	−	0.0029 ± 0.0019	16.7 ± 0.4	0.006 ± 0.002

[a] From stopped-flow or manual folding experiments at 20 °C in 0.087 M GuHCl/10 mM phosphate buffer at pH 7.0. Amplitudes are normalized relative to the total signal change between fully unfolded and folded states. [b] Final Trx concentration of 20 μM. [c] Final Trx concentration of 70 μM. [d] Final Trx concentration of 4 μM.

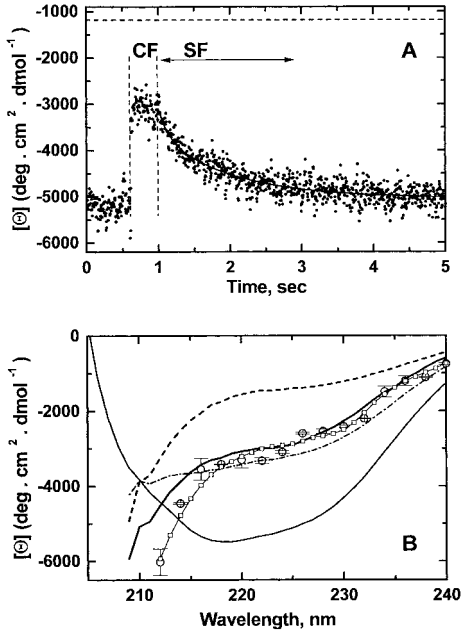


FIGURE 2: (A) Far-UV CD spectrum of the Trx refolding burst phase. Continuous- and stopped-flow CD kinetic trace at 228 nm of Trx refolding at 20 °C. The continuous-flow phase (CF) was monitored during the continuous injection over the course of 400 ms of 200 μL of denatured Trx (10 mg/mL) in 4 M GuHCl, with 3.8 mL phosphate buffer. The flow was then interrupted, and data were acquired for 4 s. The CD signal sampling interval was 5 ms for all the experiments. Recording of the ellipticity was triggered before the CF phase in a noninjective 600 ms phase in which the plateau of the previous kinetic trace was recorded using the same filtering time constant (5 ms). (B) Experiments identical to that illustrated in panel A were repeated at various wavelengths. The transient CD spectrum of the burst phase was obtained by plotting the ellipticity (averaged values between the CF mode and the stopped-flow mode data obtained by back extrapolation) as a function of wavelength and is shown as large open circles (○). The dash−dot line represents the best fit for the burst phase spectrum as a linear combination between the spectra of the native and denatured proteins. The thick solid line represents the best fit for the burst phase spectrum as a linear combination between the spectrum of the denatured protein and the spectrum calculated from the known secondary structure of native Trx in the absence of any contribution from the side chains. The small open squares (□) represent the deconvoluted burst spectrum using Contin [four denatured peptides were added to the original reference set, according to Venyaminov et al. (*43*)]. The light solid and dashed lines correspond to the spectra of refolded (in 0.2 M GuHCl) and unfolded (in 4 M GuHCl) thioredoxin, respectively, at equilibrium in the corresponding buffer and at the same protein concentration, obtained by scanning at equilibrium in the stopped-flow cell.

deduced from X-ray crystallography data (*11*). Though a quantitative analysis of far-UV CD spectra can never be considered entirely reliable, in particular for the prediction of β-structure content, the large deficiency in α-helix and excess in β-structure compared to those for native Trx suggested two possible interpretations. The burst intermediate might be a homogeneous population of molecules with an incomplete or incorrect secondary structure. Or it might contain a mixture of molecules with native-like secondary structure (10−25% by comparing the α-helix prediction shown in Table 2 with the native α-helix content) and of molecules (75−90%) with only random and β-structure. The latter population would then have a β-structure content of about 33−35%, a value not incompatible with that of native Trx.

Table 2: Deconvolutions of the Far-UV CD Spectrum of the Burst Intermediate Formed during the Refolding of Denatured Oxidized *E. coli* Thioredoxin

method	% α-helix	% β-structure	% turns and other	total %	rms
K2D[a]	6	39	56	101	0.182
Contin[b]	6	33	54	93	0.143
Selcon[c]	10	35	55	100	0.378
Prosec[d]	4	36	47	87	0.205
native Trx (X-ray data[e])	40	30	30	100	−

[a] Andrade et al. (*26*). [b] Provencher and Glöckner (*24*) and Venyaminov et al. (*25*). [c] Sreerama and Woody (*23*). [d] Yang et al. (*27*). [e] Katty et al. (*11*).

Stability of the Burst Intermediate. Stopped-flow Trx refolding measurements were carried out in the presence of various concentrations of GuHCl at 20 and 6 °C while recording the ellipticity at 222 nm. Refolding at different final GuHCl concentrations was also monitored by fluorescence, using either manual mixing or the stopped flow to dilute the denaturing agent. Different acquisition times were selected, depending on the final GuHCl concentrations (see Materials and Methods). The data thus obtained were analyzed by nonlinear least-squares fitting methods using a two-exponential model, which provided the rate constants and amplitudes of the very rapid and of the rapid phases, as well as the amplitude of the burst phase (see Table 1). The GuHCl dependence of the amplitude of the burst and the very rapid and the rapid phases was quantitatively analyzed using the standard equation derived for a two-state transition (see Materials and Methods). The corresponding midpoints C_m and the free energies $\Delta G°$ are reported in Table 3. At 20 °C, the three transitions were cooperative with a midpoint well below the transition zone of the native protein. In the absence of a specific model for the folding pathway, the kinetic phases observed experimentally cannot be quantitatively interpreted. However, because it is very well resolved

Table 3: Stability of Intermediates Formed during the Refolding of Denatured Oxidized *E. coli* Thioredoxin

	far-UV CD				fluorescence			
	C_m (M)		$\Delta G°$ (kcal/mol)		C_m (M)		$\Delta G°$ (kcal/mol)	
phase	20 °C	6 °C	20 °C	6 °C	20 °C	6 °C	20 °C	6 °C
BP	0.6 ± 0.3	*b*	3.5 ± 1.8	1.3 ± 0.6	1.3 ± 0.3	*b*	4.3 ± 1.1	1.7 ± 1.0
VRP	1.3 ± 0.2	1.0 ± 0.2	4.6 ± 0.8	3.9 ± 1.1	1.5 ± 0.3	0.9 ± 0.3	6.3 ± 1.2	5.2 ± 0.8
RP	1.7 ± 0.2	*c*	5.1 ± 1.1	*c*	1.6 ± 0.2	1.0 ± 0.3	6.3 ± 1.4	2.2 ± 0.5
MP[a]	1.5	na	na	na	1.5	na	na	na
SP[a]	1.7	na	na	na	1.7	na	na	na
N ↔ D (eq)	2.7 ± 0.2	nd	2.6 ± 0.1	nd	9.6 ± 1.0	nd	9.5 ± 1.0	nd

[a] From Kelley *et al* (*13*); *n.a.*, data not available. [b] Not determined because of the absence of a native baseline. [c] Not determined because of an insufficient number of data points.

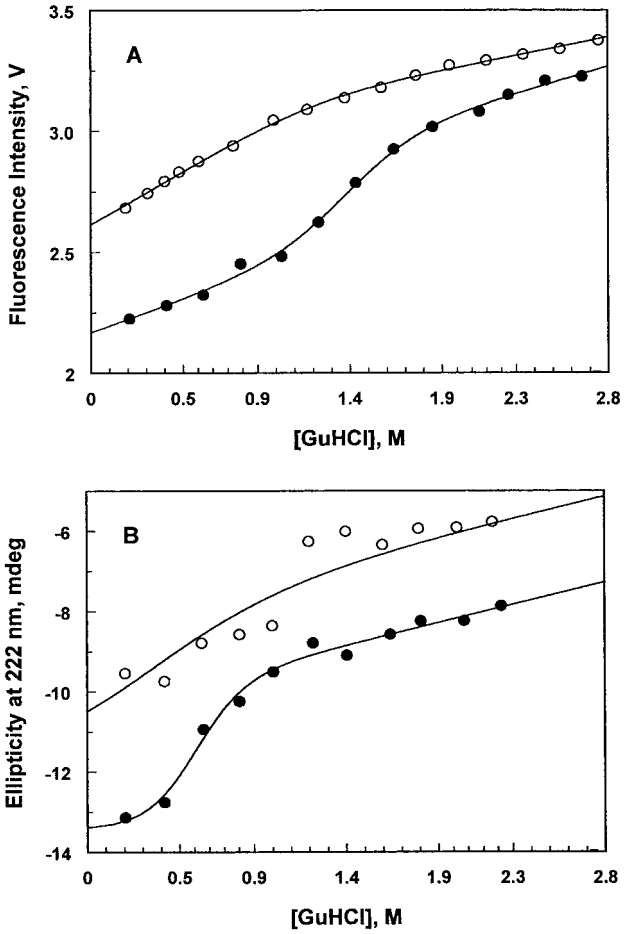


FIGURE 3: Stability of the burst phase. The fluorescence (A) or ellipticity at 222 nm (B) of the burst phase was determined by back extrapolation to time zero (see Figure 2) of the kinetic traces obtained in experiments performed at 6 (○) and 20 °C (●) and at various GuHCl concentrations. The amplitudes are represented as a function of GuHCl concentration. The solid curves are fits based on a two-state model with linear dependence of the unfolding free energy on denaturant concentration (eq 2).

in time, the amplitude of the burst phase does describe the concentration of the burst intermediate. Thus, variations of its amplitude with the residual GuHCl concentration reflect the stability of the burst population, which could be studied by following the transitions in fluorescence and far-UV CD at 20 °C (Figure 3A,B, closed symbols). That the two corresponding midpoints were different (0.7 M GuHCl with far-UV CD and 1.3 M GuHCl with fluorescence) suggested that the population present at the end of the burst phase might be heterogeneous. At 6 °C, neither fluorescence nor far-UV CD showed a clear cooperative GuHCl-induced transition (Figure 3A,B, open symbols).

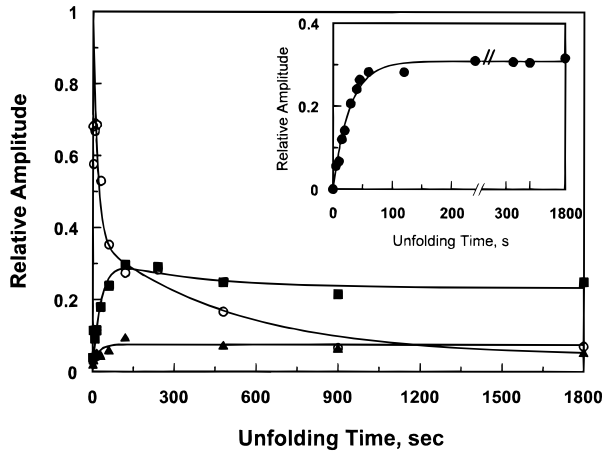


FIGURE 4: Stopped-flow and manual multimixing experiments on thioredoxin at 0.315 M residual GuHCl. The amplitudes of the VR phase (○), R phase (■), and M phase (▲) detected by fluorescence are normalized and plotted vs unfolding time. The inset shows the variation of the relative amplitude for the slow refolding phase with the unfolding time (●).

Double-Jump Experiments. Previous results with a Pro 76 variant indicated that a slow isomerization that occurs in the unfolded state was responsible for the slow phase of Trx refolding (*15*). To investigate whether proline isomerizations might also be involved in the faster phases of Trx folding, manual and stopped-flow double-jump experiments were performed as follows. Native Trx at 10 mg/mL in KP_i buffer supplemented with 1.3 M sucrose (for density compensation and minimization of remixing artifacts in the stopped flow) was diluted 10-fold with 7 M GuHCl to give a final, strongly denaturing concentration of 6.3 M GuHCl (unfolding half-time ∼ 10 ms). After incubation for a timed interval (1 s to 30 min) in the denaturing buffer, the unfolded protein was diluted 20-fold with KP_i buffer (final GuHCl concentration of 0.315 M) and the protein fluorescence was recorded along the refolding process. The initial observations of Kelley and Stellwagen (*14*) on the slow phase were confirmed by manual mixing experiments (inset of Figure 4). The effects of the time of unfolding on the relative amplitudes of the very rapid, rapid, and medium folding phases are shown in Figure 4. The amplitude of the very rapid phase decreased, while the amplitudes of the rapid and medium phases increased with increasing unfolding time. Interestingly, the amplitude of the very rapid and rapid phases varied with the unfolding time according to a biphasic kinetic process with the same two rate constants $k_{obs,1}$ of 0.036 ± 0.003 s^{-1} and $k_{obs,2}$ of

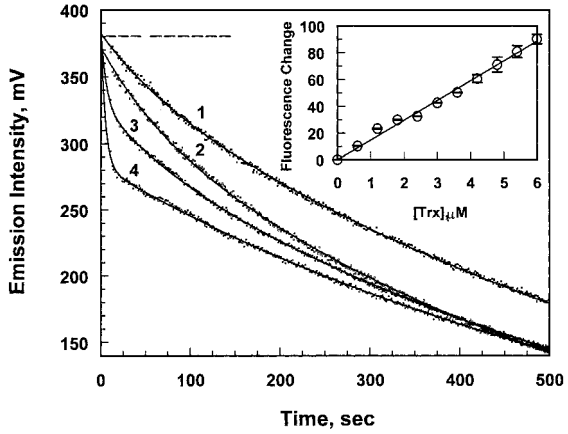


FIGURE 5: Reactivation kinetic measurements of *E. coli* thioredoxin. Kinetic measurements of the oxidation of NADPH by native thioredoxin were performed by recording the fluorescence at 450 nm, under the experimental conditions described in Materials and Methods (trace 4). The reactivation kinetic profiles of Trx denatured in 4 M GuHCl over the course of 10 s (trace 3) or 3 h (trace 2) were obtained by manual multimixing experiments, keeping constant the total volume and the concentrations of the different reagents. Controls with NADPH alone (dashed line) or NADPH in the presence of thioredoxin reductase only (trace 1) were recorded under the same experimental conditions. The inset shows the amplitude of the rapid NADPH oxidation phase as a function of the amount of native Trx in the assay.

$2.4 \times 10^{-3} \pm 1.5 \times 10^{-4}$ s^{-1}. The fit of the relative amplitude variation of the medium phase (displaying the same value for $k_{obs,1}$) is not incompatible with a two-exponential process since, due to the lack of precision in the measurements for such a small range of amplitude variation, it is difficult to ascertain whether a second phase exists. For an unfolding time of 1800 s, the relative amplitudes of all four phases were consistent with those determined previously for the corresponding phases in the whole refolding process (24 h denaturation in 4 M GuHCl).

Activity Measurements. We first attempted to use the classical coupled assay of Trx based on the reduction of DTNB in the presence of thioredoxin reductase (*20*) to monitor the kinetics of appearance of the fully native protein during a refolding experiment. However, while the steady-state kinetic parameters determined under the usual conditions ($k_{cat} = 25 \pm 1$ s^{-1} and $K_M = 2.54 \pm 0.19$ μM) were consistent with previous measurements (*32*), this assay could not be used under the conditions of our refolding experiments (i.e., in the presence of GuHCl at $\geq$87 mM) because under these conditions thioredoxin reductase reduces DTNB very rapidly even in the absence of thioredoxin. We therefore developed an "active site titration" assay (see Materials and Methods) based on the idea that one molecule of native oxidized Trx will be specifically and rapidly reduced by one molecule of thioredoxin reductase which, in turn, will be reduced by one molecule of NADPH. Thus, one should observe, by monitoring the fluorescence of NADPH (excitation at 350 nm and emission at 450 nm) a 1:1 stoichiometric reduction of native Trx by NADPH. Figure 5 shows such a titration experiment, performed in the presence of 87 mM GuHCl with 6 μM native Trx. The trace obtained (Figure 5, trace 4) showed two phases, a rapid one completed in less than 20 s with a rate constant of 0.1 s^{-1} and a slow one extending well over 10 min. The latter one is clearly due to the oxidation of NADPH (presumably by the oxygen of air) in the presence of thioredoxin reductase, as indicated in a control without Trx (Figure 5, trace 1) which shows this slow phase, but not the rapid one. The amplitude of the rapid phase was proportional to the amount of native Trx in the assay (inset of Figure 5). When Trx, unfolded for 3 h or by overnight incubation in 4 M GuHCl, was diluted in the assay mixture to 6 μM (87 mM residual GuHCl), the kinetics of NADPH oxidation (Figure 5, trace 2) were faster than those in the control without Trx (trace 1), but much slower than those with native Trx (trace 4). That traces 2 and 4 finally merged after about 500 s of refolding is consistent with the observation that, at this time, the folding of Trx was essentially completed (*28*) and hence the protein was entirely reduced. Thus, the delay of trace 2 as compared to trace 4 reflected the kinetics of the Trx refolding process. Extrapolation of trace 2 to the end of the hand-mixing dead time (6−11 s; see Materials and Methods) showed the existence during the dead time of a fluorescence decrease corresponding to about 5−8% of the rapid decay observed with native Trx (trace 4). To determine whether the small rapid fluorescence decrease observed with unfolded Trx reflects the rapid folding of a small fraction of molecules into the native conformation, manual double-jump reactivation experiments were performed. The refolding conditions, in particular, the final concentrations of thioredoxin, NADPH, thioredoxin reductase, and GuHCl, were the same as above (see Materials and Methods). But the time of unfolding of Trx in concentrated GuHCl was varied. For short unfolding times of 10 or 15 s, a rapid phase with a large amplitude was observed for the NADPH fluorescence decay (Figure 5, trace 3). Its rate constant (0.15 s^{-1}) was very close to that observed with native Trx (0.1 s^{-1}; see above), and its amplitude was about 60 and 50%, respectively, of that observed for the rapid phase of NADPH fluorescence decay observed when titrating native Trx. These observations indicate that the rapid phase of NADPH fluorescence decay observed in double-jump experiments indeed corresponds to the rapid titration of native Trx molecules. That both the amount of these rapidly titratable species and the amount of very rapidly folding intermediate (see Figure 4) decrease in parallel with the time of unfolding in double-jump experiments strongly suggests that the very rapid phase observed by far-UV CD and fluorescence during the refolding of Trx reflects the direct formation of native, active Trx from the burst intermediate. That the population of very rapidly refolding molecules decreases with increasing unfolding times indicates that the other populations of unfolded Trx molecules increase as a result of "slow" isomerizations, possibly involving prolyl residues, that take place in the unfolded state.

Temperature Dependence of the Folding Rate Constants. As a further test of a possible role of proline isomerizations in the various phases that were observed during the refolding process, the effect of temperature on the rate constants was investigated. For this purpose, kinetic measurements were performed in the presence of 0.2 M GuHCl at different temperatures ranging from 6 to 30 °C, using both manual and stopped-flow mixing. The rate constant for each of the refolding phases was shown to be strongly dependent on temperature. For each phase, the Arrhenius plot showed a linear dependence of the logarithm of the rate constant on the reciprocal of the temperature (data not shown). The

activation energies of the very rapid, rapid, medium, and slow refolding phases, derived from the slopes of the Arrhenius plots, were estimated to be 12.3 ± 1.7, 15.2 ± 1.8, 13.4 ± 0.4, and 17 ± 2 kcal/mol, respectively. These values, though not very precise, are compatible with those reported for the activation energies of proline (trans−cis) isomerization: 15 (*33*) and 18 kcal/mol (*34*).

DISCUSSION

In this report, we have described the kinetics of refolding of Trx as monitored by five independent criteria: the far-UV CD which mainly reflects the formation of secondary structure, the binding of ANS which allows detection of the presence of a loosely packed hydrophobic core, the near-UV CD which depends on the tight packing and immobilization of aromatic and cystinyl side chains, the intrinsic fluorescence which characterizes the polarity of the environment of the Trx tryptophanyl residues that are both located close to the active site, and the activity which is used to monitor the complete renaturation of Trx. Using these signals, we could identify and characterize five phases during the refolding of Trx. Their kinetic properties will first be discussed, together with the structural changes associated with each phase.

Burst Phase (<4 ms). The first evidence (data not shown) indicating that an early intermediate was formed in a burst phase was brought to light by the observation that, at GuHCl concentrations below 1.5 M, the logarithm of the apparent rate constant of the first (the "very rapid") observable kinetic phase did not vary linearly with the GuHCl concentration. This already suggested that under these conditions an intermediate state had to be formed during the dead time before the very rapid phase could occur. A more direct demonstration of the formation of a burst intermediate was provided by the strong decrease in the fluorescence and far-UV CD signals that occur, at GuHCl concentrations below 1.5 M, during the 3−4 ms dead time of the stopped-flow measurements. Thus, at 0.087 M GuHCl, for example, nearly 68% of the total ellipticity and 30% of the total fluorescence change associated with refolding occur during the dead time (Table 1). In terms of conformation, the Trx molecules present at the end of the burst phase can be described as follows. The far-UV CD spectrum unambiguously demonstrates the presence of important amounts of repetitive secondary structure (about 40−45% in 0.2 M residual GuHCl; see Table 2). Though the transient far-UV CD spectrum of the burst intermediate does not permit a precise deconvolution (mainly because of the rather limited wavelength range investigated), and though its quantitative analysis (see Table 2) should therefore be considered with much caution, this spectrum clearly shows that β-structures have already been building up during the burst phase. Although a significant fraction of the far-UV ellipticity may well arise from non-native secondary structures formed during the burst (*6*, *35*, *36*), the fact that mostly β-structures are represented at this stage leads to the hypothesis that the folding of Trx might start with the rapid formation of the β-sheet core of the protein.

The lack of a near-UV CD signal in this burst intermediate implies that its core is still a long way from being tightly packed. Yet, the 30% decrease in fluorescence intensity of the burst intermediate relative to the unfolded state indicates that the polypeptide chain may have either undergone a hydrophobic collapse or formed some local hydrophobic clusters. Finally, the burst intermediate contains solvent accessible hydrophobic areas, as indicated by its ability to bind ANS (Table 1), that may contribute to its stability. As expected, this dye binding capacity tends to be maximal at an early stage and to be lost quite rapidly thereafter, confirming that these highly transient species are fundamentally different from the native molecules in their level of structural organization (*7*, *37*). Taken together, these observations strongly suggest that hydrophobic interactions play a key role in the formation of this very early intermediate, as postulated in the hydrophobic collapse or molten globule models.

Our observations on the effects of GuHCl and temperature on the stability of the burst intermediate have confirmed this role of hydrophobic interactions. Indeed, the burst amplitude of the far-UV CD at 20 °C shows a clear GuHCl-induced unfolding transition and vanishes at GuHCl concentrations above 1.5 M, i.e., well below the transition midpoint of native Trx (Figure 3). In contrast, the GuHCl dependence of the burst phase amplitude at 6 °C showed no indication of a clear GuHCl transition, and a partial "unfolding" of the burst intermediate seems to already occur at that temperature even in the absence of GuHCl. This destabilization of the intermediate at low temperatures, shown by the increased sensitivity to GuHCl of both its far-UV CD and fluorescence, strongly suggests that the stability of the burst intermediate primarily depends on hydrophobic interactions, which would indeed be destabilized at low temperatures.

The stability experiments in GuHCl performed at 20 °C show distinct transition curves for the burst intermediate; the transition midpoint is at 0.6 M for the far-UV CD and at 1.3 M for the fluorescence. Such differences between the two measurements indicate a deviation from a two-state transition, which could be expected due to the particular positions of the two tryptophan residues present in Trx. Indeed, while the far-UV CD reflects the overall conformation of the polypeptide chain, the two tryptophan residues that afford most of the fluorescence signal are very near each other and close to a protrusion likely to behave independently of the rest of the molecule.

As discussed above, the burst intermediate has significant secondary structure and no fixed tertiary structure, hydrophobic interactions are involved in its stabilization, and it contains accessible hydrophobic clusters, typical of a molten globule. Yet, the fact that the burst intermediate is not a homogeneous species and contains some non-native secondary structure elements (i.e., secondary structure elements that do not exist in native Trx; see Far-UV CD Spectrum of the Burst) leads us to believe that it cannot be described as a real molten globule which, according to the original definition, should have only native-like secondary and supersecondary structures. Rather, it could be described well as a pre-molten globule state (*35*), a fluctuating ensemble containing many conformers in rapid equilibrium in which secondary structure elements, both native and non-native, are not yet stabilized by long-range contacts. One can therefore predict that, as for several other proteins (*38*, *39*), the burst intermediate of Trx would have amide protons poorly protected from exchange with the solvent.

Very Rapid Phase. The next phase we observed following the burst was characterized by a rate constant of about 33 ± 3 s^{-1} at 0.087 M GuHCl. At 0.2 M residual GuHCl (where measurements were more precise), the rate constant was 6.5 ± 0.9 s^{-1}, and the amplitude of this phase, which indicated an additional regain of secondary structure, corresponded to a further increase of about 30% of the amplitude of the far-UV ellipticity as compared to the burst. The transient far-UV CD spectrum obtained at the end of this phase (data not shown) was consistent with a contribution of about 80% of a native-like CD spectrum. Stability experiments monitored by fluorescence and far-UV CD showed similar midpoint concentrations ($C_m \sim 1.3-1.5$ M) for the GuHCl-induced transitions at 20 °C, with only very small differences in cooperativity. At low temperatures, the titration curves in fluorescence and far-UV CD still showed a clear transition pattern, but were displaced to lower GuHCl concentration midpoints (0.9−1 M) (data not shown). The fact that this phase exhibited a signal in near-UV CD with an amplitude of about 20% indicates that, during this very rapid phase, some aromatic side chains in the core start gaining more defined conformational restrictions. Yet, the polarity of the environment of the tryptophan residues does not change much, as indicated by the small additional fluorescence quenching of only 7% compared to the total amplitude change occurring upon complete refolding of Trx.

Rapid Phase. During this phase, characterized by a relaxation time of 1.58 ± 0.06 s, the transient far-UV CD spectrum displayed an additional increase of 12% in the amplitude of the far-UV ellipticity. Also on this time scale, the largest changes in the near-UV CD and fluorescence amplitude occurred: 45 and 40%, respectively (see Table 1). Stability experiments using far-UV CD and fluorescence again showed the same transition midpoints centered around 1.6 M GuHCl, indicating that the intermediates formed during the rapid phase gained some cooperativity. This, together with the important changes in the near-UV CD, might reflect the formation of some long-range tertiary interactions.

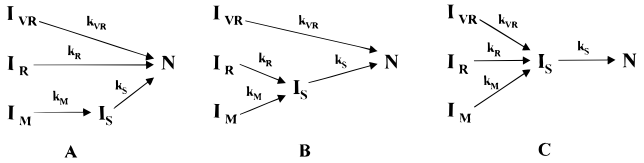
Medium and Slow Phases. The medium phase was characterized by a very small change in the amplitude of the far-UV CD signal. Similarly, the fluorescence intensity changed very little (only 6% of the total amplitude). The most important change that occurred during this phase was in the near-UV CD, which reached its final plateau through an increase of 36% of the total renaturation signal. Thus, the medium phase primarily involves side chain packing. The last, slow phase was characterized by a decrease in the amplitude of the far-UV CD ellipticity (−5.6%) and a significantly larger change in fluorescence (about 17%). Though the rate constant of the slow activity regain (Figure 5, trace 2) could not be determined with precision, this slow regain clearly occurred in the same time range as the slow phase observed by CD and fluorescence, and probably corresponds to the same folding step. This conclusion is supported by the following considerations. The slow phase of CD and fluorescence changes has been attributed in previous papers (*14*, *15*) to a trans−cis isomerization involving Pro 76. The relative proportions of rapid and slow phases of activity regain we observed upon refolding of fully unfolded Trx (about 10 and 90%, respectively) are similar to those of cis and trans isomers of Pro 76, respectively, in denatured Trx estimated from studies of Trx fragments (*28*). Thus, the trans−cis isomerization of Pro 76 also seems to be responsible for the slow activity regain during Trx refolding. The relatively large fluorescence quenching (17% of the total change) that appears during this phase may be explained by the fact that, in native Trx, Pro 76 is in the protrusion of the active site near the two tryptophan residues. Thus, the last event occurring on the refolding pathway seems to be the final organization of the active site protrusion itself.

General Refolding Mechanism and Comparison with Previous Kinetic Studies. Previous experiments (*13*−*15*) detected only three phases during the refolding of *E. coli* thioredoxin. This was presumably due to the choice of experimental conditions, since the refolding was studied in the presence of 2 M GuHCl, a concentration at which the two first intermediates we identified are strongly destabilized. It was also presumably due to limitations in the instrumentation, which did not permit large enough variations in the dilution factors. The kinetic measurements reported here were conducted over a range of dilution factors and of residual GuHCl concentrations, allowing resolution of at least five phases during the refolding of Trx. The double-jump experiments we report, performed by both manual and stopped-flow mixing, demonstrated that the amplitudes of all the four phases that could be studied (very rapid to slow) were controlled by slow isomerization reactions occurring under the unfolding conditions that lead to a partitioning of the denatured state into subpopulations of unfolded molecules in slow equilibrium that give rise to each of the observable kinetic phases. That equilibrium between these subpopulations is ultimately achieved is confirmed by our observation that after 30 min of unfolding in double-jump experiments, the amplitude of each phase reached a finite plateau (Figure 4) that corresponded closely to the amplitude determined when observing the refolding of Trx incubated overnight in 4 M GuHCl (Table 1). The kinetics of the isomerizations that occur in the unfolded state are reflected by the kinetics of variation with the unfolding time of the amplitudes of the various phases. Thus, the proportion of very rapidly folding molecules, reflected by the amplitude of the very fast refolding reaction, appears to decrease with time of unfolding to about 6% of its initial value, according to a biphasic kinetic process with time constants τ_{vr1} and τ_{vr2} of 33 and 400 s, respectively.

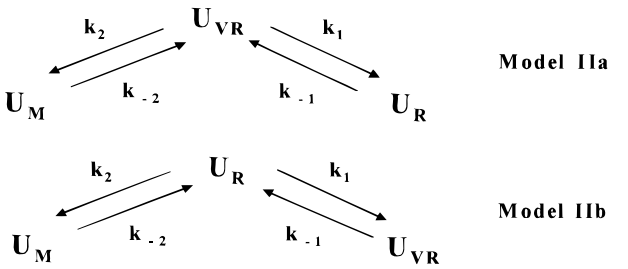
Our results are not consistent with the previously proposed "cubic" model of Trx refolding (*29*, *30*). Indeed, a strong prediction of the cubic model is that the kinetics (as a function of unfolding time) of the amplitude change in stopped-flow double-jump measurements should show three relaxation times, which can be observed in at least one of the refolding phases. Yet, none of the refolding phases we observed showed amplitude dependence (with respect to unfolding time) with more than two phases. Moreover, while the previous models suggested that some unfolded Trx molecules refold directly into the native state, we failed to detect any molecule that does not fold via the burst species. These considerations prompted us to develop a more appropriate folding model.

This model must take into account the mass conservation equation. Because two, and only two, relaxations were observed in the unfolded state, the model assumes that the denatured state should be populated with three, and no more

Scheme 1



Scheme 2: Equilibria in the Denatured State



than three, subpopulations in equilibrium. Finally, any model must take into account the observed burst phase, leading to an apparently heterogeneous population that is followed by four refolding phases. With these considerations in mind, three models describing the Trx refolding pathway could be imagined.

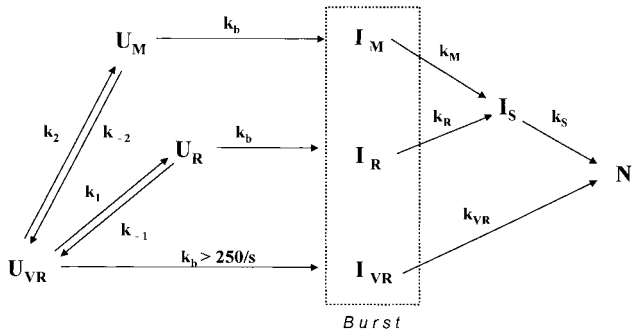
These models are depicted in Scheme 1, in which it is assumed that each of the three subpopulations present in the unfolded state, called U_{VR}, U_R, and U_M, generates the corresponding burst intermediate I_{VR}, I_R, and I_M extremely rapidly (less than 4 ms). Some of these intermediate species are converted directly into native Trx, while the others must first isomerize to I_S before folding during the slow phase into the native conformation. In all three cases, I_S cannot be generated directly from a fourth distinct subpopulation that would be present in the denatured state since, as stated above, the denatured state can contain only three subpopulations. Rather, I_S must result from the further folding of at least one of the burst intermediates.

According to Scheme 1, the amplitude of each refolding phase should depend on the concentration of the burst intermediate(s) on the corresponding pathway, and hence on the concentration of each subpopulation in the unfolded state. Accordingly, the kinetics of variation of the amplitude of each refolding phase in a double-jump experiment, and hence the corresponding apparent rate constants, should reflect the relaxation of the equilibria between these subpopulations under the unfolding conditions. In fact, these apparent rate constants should result from a mixture of the rate constants k_1, k_2, k_{-1}, and k_{-2} (Scheme 2) according to eq 3:

$$k_{obs,1,2} = -^1/_2[(k_1 + k_{-1} + k_2 + k_{-2}) \pm \sqrt{[(k_1 + k_{-1}) - (k_2 + k_{-2})]^2 + 4k_{-1}k_2}] \quad (3)$$

From this equation and with the values of the relative amplitudes at infinite unfolding time for each phase (Table 1), we could estimate for each of the two possible equilibria depicted in Scheme 2 the values of k_1, k_2, k_{-1}, and k_{-2} and simulate the unfolding time dependence of the relative amplitudes of the very rapid, rapid, and medium refolding phases (Supporting Information), using the VisSim software package (Visual Solution Inc.). The simulated kinetics, which for these three phases do not depend on the pathway

Scheme 3



chosen among those represented in Scheme 1, were compared with the experimental data (Figure 4) obtained in stopped-flow double-jump experiments. Simulations using equilibrium IIb showed a lag period for the variation of the medium phase amplitude, which was not observed experimentally (Figure 4). Therefore, equilibrium IIb was rejected, and equilibrium IIa was retained to describe the unfolded state. Thus, this equilibrium controls the relative concentrations of the three burst intermediates I_{VR}, I_R, and I_M of the models depicted in Scheme 1. Which of these three models is most consistent with the folding kinetics? Model Ic, where all molecules refold via I_S, predicts that the regain of activity should show only a slow phase. That a significant fraction of the unfolded protein refolds very rapidly to active Trx in double-jump experiments therefore rules out model Ic. The precision of the activity measurements in the double-jump experiments did not allow us to determine the kinetics of decay of the rapidly renaturable species with unfolding time. Yet, the fractions of rapidly renaturable protein present after 10 or 15 s of unfolding (60 and 50%, respectively) and after an infinite unfolding time (about 6% only) are comparable with the amplitudes observed at these unfolding times for the very rapid phase detected in fluorescence (Figure 4 and Table 1). Thus, the decay of the rapidly renaturable species parallels the decay of the amplitude of the very rapid refolding phase, indicating that I_{VR} and only I_{VR} directly leads to native Trx, which also rules out model Ia. The latter conclusion should however be seen as tentative only because the relative amplitudes of the four phases very much differ depending on the signal used (Table 1), and therefore cannot be considered to represent the concentrations of the intermediate species. Consequently, though model Ia can by no means be definitively discarded at this stage, we shall tentatively privilege model Ib in further discussions. This conclusion is incorporated in Scheme 3 below, which shows the model we propose to describe the folding of Trx.

This model, as well as the two related ones described in Scheme 1, assumes that each of the species present in the unfolded state (U_{VR}, U_R, and U_M) can form burst intermediates (I_{VR}, I_R, and I_M) containing different amounts of ordered structure, each one able to fold at different rates either into the native conformation or into the slowly folding intermediate I_S. From the double-jump experiments as well as from the value of the activation energy we determined for the slow phase [17 kcal/mol, a value close to that reported by Lin and Brandts (*34*) for proline isomerization], it seems likely that this phase is controlled by a proline isomerization. This is compatible with previous experiments with the P76A variant and the wild type protein (*15*, *40*), which led to the

conclusion that a trans−cis isomerization of Pro 76 was responsible for the slow refolding phase of Trx. Thus, I_S would have its Pro 76 in the trans conformation.

It had been previously concluded that, for Trx, only the slowest refolding kinetic phase could generate a product having the activity and stability of the native protein (*14*). This conclusion is ruled out by the double-jump reactivation measurements (Figure 5) with 10 or 15 s of unfolding, which clearly demonstrate that there are at least two distinct refolding pathways toward native Trx, one passing through I_{VR} and the other through I_S. As just discussed, the initial isomerization state of Pro 76 in the unfolded state may be one of the major differences between these pathways. Because thioredoxin contains five prolyl residues, it is likely that the distribution of cis−trans prolyl isomers in the denatured state may also be responsible for the partitioning of Trx into the three subpopulations present in the unfolded state. This is compatible with the generally accepted view that protein folding reactions, which are not coupled with proline isomerizations, are often rapid (*41*), as opposed to folding reactions controlled by prolyl isomerizations. Accordingly, in our model, U_{VR} would correspond to the U ($P76_c$, all other P_{trans} species) that contains only native cis or trans prolyl isomers. This species would first collapse, during the burst, into a pre-molten globule type population of molecules with only the correct prolyl isomers, which in turn could fold during the very rapid phase into the native protein. The products of the partial folding reactions of U_R and U_M would be folding intermediates (I_R and I_M) with Pro 76 in the non-native conformation. The structural differences existing between U_R and U_M might involve proline residues, other than Pro 76, trapped in the non-native configuration in the denatured state. One should however also consider the possibility of other isomerization reactions occurring on the same time scale. Such isomerizations might involve another cis peptide bond, at one or several nonproline residues, as already reported for slow folding events in other proteins (*42*). Though the model we propose (Scheme 3) is certainly not definitely proven by the kinetic studies reported here, it is entirely compatible with our experimental results. It is also compatible with the growing amount of experimental evidence and theoretical modeling suggesting that, if the condensation of a protein leads to a near-native arrangement of its major structural elements, subsequent folding to the native state may occur within milliseconds (*43*, *44*).

Moreover, and regardless of the model used to describe the folding of Trx, the results reported above inevitably lead to the following important conclusions. Though Trx is a small one-domain protein, it does not obey a two-state model. At least one of the intermediates we observed, that present at the end of the burst, does not result from a proline isomerization. Distinct folding pathways can lead to native Trx. The pathway followed by a given molecule is controlled by slow isomerization reactions (including at least the isomerization of Pro 76) that occur in fully unfolded Trx. These conclusions fit better with the "new view" of protein folding that has emerged during the past decade than with the purely sequential models formerly used to account for the presence of folding intermediates.

SUPPORTING INFORMATION AVAILABLE

Analysis of the kinetic model, mathematical formalism, and two figures containing the results of the simulation procedure (5 pages). Ordering information is given on any current masthead page.

REFERENCES

1. Wolynes, P. G., Onuchic, J. N., and Thirumalai, D. (1995) *Science 267*, 1619−1620.
2. Elöve, G. A., Chaffotte, A. F., Roder, H., and Goldberg, M. E. (1992) *Biochemistry 31*, 6876−6883.
3. Walkenhorst, W. F., Green, S. M., and Roder, H. (1997) *Biochemistry 36*, 5795−5805.
4. Khorasanizadeh, S., Peters, I. D., Butt, T. R., and Roder, H. (1993) *Biochemistry 32*, 7054−7063.
5. Jennings, P. A., and Wright, P. E. (1993) *Science 262*, 892−896.
6. Chaffotte, A. F., Guillou, Y., and Goldberg, M. E. (1992) *Biochemistry 31*, 9694−9703.
7. Itzhaki, L. S., Evans, P. A., Dobson, C. M., and Radford, S. E. (1994) *Biochemistry 33*, 5212−5220.
8. Kiefhaber, T., Bachmann, A., Wildeger, G., and Wagner, C. (1997) *Biochemistry 36*, 5108−5112.
9. Neira, J. L., and Rico, M. (1996) *Folding Des. 2*, R1−R11.
10. Matthews, C. R. (1993) *Annu. Rev. Biochem. 62*, 653−683.
11. Katti, S., LeMaster, D. A., and Eklund, H. (1990) *J. Mol. Biol. 212*, 167−184.
12. Dyson, H. J., Holmgren, A., and Wright, P. E. (1989) *Biochemistry 28*, 7074−7087.
13. Kelley, R. F., and Stellwagen, E. (1984) *Biochemistry 23*, 5095−5102.
14. Kelley, R. F., Wilson, J., Bryant, C., and Stellwagen, E. (1986) *Biochemistry 25*, 728−732.
15. Kelley, R. F., and Richards, F. M. (1987) *Biochemistry 26*, 6765−6774.
16. Kemple, M. D., Yuan, P., Nollet, K. E., Fuchs, J. A., Silva, N., and Prendergast, F. G. (1994) *Biophys. J. 66*, 2111−2126.
17. Schägger, H., and von Jagow, G. (1987) *Anal. Biochem. 166*, 368−379.
18. Pace, C. N. (1986) *Methods Enzymol. 131*, 266−280.
19. Holmgren, A., and Reichard, P. (1967) *Eur. J. Biochem. 2*, 187−196.
20. Luthman, M., and Holmgren, A. (1982) *Biochemistry 21*, 6628−6633.
21. Santoro, M. M., and Bolen, D. W. (1988) *Biochemistry 27*, 8063−8068.
22. Manavalan, P., and Johnson, W. C., Jr. (1987) *Anal. Biochem. 167*, 76−85.
23. Sreerama, N., and Woody, R. W. (1993) *Anal. Biochem. 209*, 32−44.
24. Provencher, S. W., and Glöckner, J. (1981) *Biochemistry 20*, 33−37.
25. Venyaminov, S. Y., Baikalov, I. A., Shen, Z. M., Wu, C. S., and Yang, J. T. (1993) *Anal. Biochem. 214*, 17−24.
26. Andrade, M. A., Chacón, P., Merelo, J. J., and Móran, F. (1993) *Protein Eng. 6*, 383−390.
27. Yang, J. T., Wu, C.-S. C., and Martinez, H. M. (1986) *Methods Enzymol. 130*, 208−269.
28. Chaffotte, A. F., Li, J.-H., Georgescu, R. E., Goldberg, M. E., and Tasayco, M. L. (1997) *Biochemistry 36*, 16040−16048.
29. Shalongo, W., Ledger, R., Jagannadham, M. V., and Stellwagen, E. (1987) *Biochemistry 26*, 3135−3141.
30. Shalongo, W., Jagannadham, M., and Stellwagen, E. (1993) *Biopolymers 33*, 903−913.
31. Ptitsyn, O. B. (1995) *Adv. Protein Chem. 47*, 83−229.
32. Navarro, J. A., Gleason, F. K., Cusanovich, M. A., Fuchs, J. A., Meyer, T. E., and Tollin, G. (1991) *Biochemistry 30*, 2192−2195.
33. Schultz, G. D., and Schirmer, R. H. (1978) *Principles of Protein Structure*, p 25, Springer-Verlag, New York.

34. Lin, D.-N., and Brandts, J. F. (1983) *Biochemistry 22*, 553−559.
35. Chaffotte, A. F., Guijarro, I., Guillou, Y., Delepierre, M., and Goldberg, M. E. (1997) *J. Protein Chem. 16*, 433−439.
36. Guijarro, J. I., Jackson, M., Chaffotte, A. F., Delepierre, M., Mantsch, H. H., and Goldberg, M. E. (1995) *Biochemistry 34*, 2998−3008.
37. Engelhard, M., and Evans, P. A. (1995) *Protein Sci. 4*, 1553−1562.
38. Radford, S. E., Dobson, C. M., and Evans, P. A. (1992) *Nature 358*, 302−307.
39. Varley, P., Gronenborn, A. M., Christensen, H., Wingfield, P. T., Pain, R. H., and Clore, G. M. (1993) *Science 260*, 1110−1113.
40. Langsetmo, K., Fuchs, J., and Woodward, C. (1989) *Biochemistry 28*, 3211−3220.
41. Scalley, M. L., Yi, Q., Gu, H., McCormack, A., Yates, J. R., and Baker, D. (1997) *Biochemistry 36*, 3373−3382.
42. Odefey, C., Mayr, L. M., and Schmid, F. X. (1995) *J. Mol. Biol. 245*, 69−78.
43. Viguera, A. R., Martinez, J. C., Filimonov, V. V., Mateo, P. L., and Serrano, L. (1994) *Biochemistry 33*, 2142−2150.
44. Villegas, V., Azuaga, A., Catasus, L., Reveter, D., Mateo, P. L., Aviles, F. X., and Serrano, L. (1995) *Biochemistry 34*, 15105−15110.

BI9805083