Articles

Recognition between Disordered States: Kinetics of the Self-Assembly of Thioredoxin Fragments†

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

Unité de Biochimie Cellulaire, Institut Pasteur, Paris, France, and Department of Chemistry, City College of the City University of New York, New York, New York 10031

Received April 11, 1997; Revised Manuscript Received October 15, 1997⊗

ABSTRACT: The disordered N- (1−73) and C- (74−108) fragments of oxidized *Escherichia coli* thioredoxin (Trx) reconstitute the native structure upon association [Tasayco, M. L., & Chao, K. (1995) *Proteins: Struct., Funct., Genet. 22*, 41−44]. Kinetic measurements of the formation of the complex (1−73/74−108) at 20 °C under apparent pseudo-first-order conditions using stopped-flow far-UV CD and fluorescence spectroscopies indicate association coupled to folding, an apparent rate constant of association [k_{on} = (1330 ± 54) M^{-1} s^{-1}], and two apparent unimolecular rate constants [k_1 = (0.037 ± 0.007) s^{-1} and k_2 = (0.0020 ± 0.0005) s^{-1}]. The refolding kinetics of the GuHCl denatured Trx shows the same two slowest rate constants. An excess of N- over C-fragment decreases the k_{on}, and the slowest phase disappears when a P76A variant is used. Stopped-flow fluorescence measurements at 20 °C indicate a GuHCl-dependent biphasic dissociation/unfolding process of the complex, where the slowest phase corresponds to 90% of the total. Their rate constants, extrapolated to zero denaturant, $k_{-1} = (9 \pm 3) \times 10^{-5}$ s^{-1} and $k_{-2} = (3.4 \pm 1.2) \times 10^{-5}$ s^{-1}, show $m^{\#}$ values of (4.0 ± 0.4) kcal mol^{-1} M^{-1} and (3.5 ± 0.1) kcal mol^{-1} M^{-1}, respectively. Our results indicate that: (i) a compact intermediate with trans P76 and defined tertiary structure seems to participate in both the folding and unfolding processes; (ii) not all the N-fragment is competent to associate with the C-fragment; (iii) conversion to an association competent form occurs apparently on the time scale of P76 isomerization; and (iv) the P76A variation does not alter the association competency of the C-fragment, but it permits its association with "noncompetent" forms of the N-fragment.

Protein recognition is central to most biological phenomena, whether they involve protein−nucleic acids or protein−protein interactions. Despite its importance, however, the interplay between structure and recognition is not well understood. Complementary protein fragments (Taniuchi & Anfinsen, 1971) that reassemble into native-like small structures can provide excellent models to study this interplay. Some of them have been tested at the atomic level (Kim et al., 1992; Sancho & Fersht, 1992; Prat Gay & Fersht, 1994; Tasayco & Carey, 1992; Tasayco & Chao, 1995). In some cases, the isolated fragments show certain ability to fold by themselves in either a native-like (Waltho et al., 1993) or a non-native-like structure (Chaffotte et al., 1991). In other cases the isolated fragments are just highly disordered (Lattman et al., 1994; Shortle, 1996a,b; Freund et al., 1996). The way the nature of the fragments affects their reassembly into the native cleaved structure can be diagnostic of the significance of local nucleation sites and/or nonlocal interactions in the process of protein recognition. Therefore, the structural analysis of isolated protein fragments and their complexes should be useful to unravel the mechanism of protein−protein recognition.

Kinetic studies of noncovalent complexes of fragments from small proteins (Prat Gay et al., 1994; Labhardt et al., 1983; Kippen & Fersht, 1995) indicate that the assembly of α-helices with β-strands can be formed with rate constants between 3−6 orders of magnitude lower than a diffusion controlled rate of collision, 10^9 M^{-1} s^{-1}. In all these cases, it appears that rate constants of association between 10^4 and 10^5 M^{-1} s^{-1} correlate with disordered protein fragments, while those showing a substantial degree of folding have either a faster or a slower rate constant, apparently depending on the competency of their folded conformations. The same behavior appears to happen during the refolding of small homodimers (Gittelman & Matthews, 1990; Milla & Sauer, 1994; Mann & Matthews, 1993; Wendt et al., 1995; Zitzewitz et al., 1995; Sosnick et al., 1996). The rate of association approaches the diffusion limit when the monomers appear to have substantial competent folded conformations despite the complexity of the final structure, while the rates are 2−3 orders of magnitude lower when the monomers appear more disordered. On the other hand, the dissociation rate constants of these homodimers and the above described cleaved proteins fall within a wide range (10^{-4}−10 s^{-1}) which seems to be related to the complexity of the assembly and the hydrophobicity of the interface between monomers. The combination of detailed structural and kinetic studies of

† This work was supported by MCB-9507255 and INT-9600006 from NSF to M.L.T., who is an NSF Career Awardee, the NIH Grant RCMI to CCNY, and awards from the Institut Pasteur CNRS (URA1128) and University of Paris VII to M.E.G.

* Author to whom correspondence should be addressed.
‡ Unité de Biochimie Cellulaire, Institut Pasteur, France.
§ Department of Chemistry, City College of New York.

⊗ Abstract published in *Advance ACS Abstracts,* December 1, 1997.

S0006-2960(97)00850-7 CCC: $14.00

association/folding processes (Zetina & Goldberg, 1982; Labhardt, 1984; Clark et al., 1997) thus appears to be a powerful tool to study protein−protein recognition. A similar approach has been taken to unravel the protein folding pathway by some groups, which either have cleaved small proteins (Sancho & Fersht, 1992; Neira et al., 1996) or have linked the monomers of a small dimer (Robinson & Sauer, 1996).

We have selected to study the association/folding process between the complementary N- (1−73) and C-fragment (74−108) of oxidized *E. Coli* Trx.[1] Each fragment comprises the sequence of at least two β-strands and an α-helix from the $\alpha\beta$-structure of the native protein (Dyson et al., 1989; Katti et al., 1990), but they appear disordered in isolation (Reutimann et al., 1981, 1983). Structural studies of the equimolar mixture of the N- and C-fragment using NMR spectroscopy indicate the presence of a noncovalent complex (1−73/74−108) with a native-like overall structural topology and subtle local conformational changes near the cleaved peptide bond 73−74 (Tasayco & Chao, 1995), which still favor the predominance of the cis P76 isomer. The striking structural similarity between this complex and its native uncleaved protein places us in an excellent position to use it as a model system to study the non-self-recognition between disordered protein fragments under nondenaturing conditions. To study the recognition process between disordered states, our first step is to compare the kinetics of formation of the same structure using two procedures: the mixing of the two disordered complementary fragments of Trx under nondenaturing conditions and the fast dilution of the chemically denatured uncleaved protein into renaturation buffer.

In this article, we report the results of studies of the association/folding process between the Trx fragments, the dissociation/unfolding of the reconstituted Trx and the refolding of the denatured uncleaved Trx using stopped-flow and manual mixing experiments together with fluorescence and CD spectroscopy. These studies focus on establishing whether a specific disordered state, or states, of one fragment is required for its recognition by the other. In addition, we test whether the isomerization of P76 is required for fragment recognition and its implications on protein folding (Dodge & Scheraga, 1996). Together, all these studies lead us to propose one of the plausible models for the recognition process involved in the reassembly of the native structure of Trx.

MATERIALS AND METHODS

Chemicals. Ultrapure GuHCl was purchased from ICN Biomedicals, Inc. All other chemicals were reagent grade. The buffer was 10 mM potassium phosphate at pH 7.0 (KP_i).

Trx Fragments. Wild-type Trx was overexpressed in *E. Coli* JF521 (gift of Dr. J. Fuchs and Dr. C. K. Woodward), purified (Langsetmo et al., 1989), and characterized by SDS−PAGE (Schagger & von Jagow, 1987). The fragments were generated (Reutimann et al., 1981; Tasayco & Chao, 1995), purified by reverse phase chromatography, and characterized by SDS−PAGE and mass spectroscopy. The C-fragment and its P76A variant were synthesized, purified and characterized using mass spectroscopy by ANASPEC (San Jose, California). The extinction coefficient of the N-fragment in KP_i (1.4×10^4 M^{-1} cm^{-1}) was measured at 280 nm (Mulvey et al., 1974). The extinction coefficient of the C-fragment in KP_i (3.9×10^4 M^{-1} cm^{-1}) was calculated at 215 nm by aminoacid analysis. The lyophilized fragments and their equimolar mixture were resuspended in 7.4 M urea/KP_i and then exchanged for KP_i. The solutions were used fresh or stored frozen at −25 °C for later use.

Stopped-Flow Fluorescence Spectroscopy. Kinetic measurements were conducted at 20 °C using a SFM-3 stopped-flow module, a monochromatic source, a fluorescence detector and software supplied by Bio/Logic (Claix-France). The sample in the flow cell (FC15, 1.5 mm, 31 μL) was irradiated with incident monochromatic light at 295 nm from a Xenon/Mercury 150 W lamp. The fluorescence emitted above 325 nm through a high-pass filter was recorded as a function of time. Controls were recorded for each fragment, the equimolar mixture, the native and unfolded Trx, the GuHCl solution and KP_i.

(a) Association/Folding. The kinetic measurements were obtained under pseudo-first-order conditions. The association was initiated by injection of equal volumes (200 μL) of each fragment in KP_i within 50 ms (3.8 ms of dead time). The kinetic traces of 10 (0−50 s) and 3 (50−1000 s) successive injections were recorded, averaged, and pasted.

(b) Dissociation/Unfolding. The denaturation of the complex was initiated by mixing 320 μL of GuHCl in KP_i (from 2.25 M to 3.5 M) with 80 μL of 40 μM complex in KP_i within 80 ms (3.8 ms of dead time). The kinetic traces of successive injections (between 4 and 10) were recorded for periods up to 500 s.

(c) Refolding of denatured Trx. The refolding was initiated by mixing 15 μL of 1.76 mg/mL previously denatured Trx, 3.5 M GuHCl in KP_i, with 585 μL of KP_i within 50 ms (2.6 ms of dead time). The kinetic traces of 50 (0−1 s), 30 (1−5 s), 15 (5−20 s), and 5 (20−1000 s) successive injections were recorded, averaged, and pasted.

Manual Fluorescence Spectroscopy. Kinetic measurements were conducted with a PTI fluorimeter equipped with an automatic shutter, a magnetic stirrer, and a water bath at 20 °C. The association was typically initiated by manually adding 200 μL of one fragment in KP_i to a 1.4 mL quartz fluorescence cell with 800 μL of the other one under vigorous stirring. The delay time between the beginning of the mixing and the acquisition was recorded for each individual kinetic trace (3−5 s). The fluorescence intensity was measured in counts per second (cps). The excitation wavelength was 295 nm with a band width of 1 nm, and the emission wavelength was 350 nm with a bandwidth of 5 nm. The kinetic traces of 3 (0−1800 s) successive additions were recorded and averaged.

CD Spectroscopy. Kinetic measurements were conducted using a spectropolarimeter CD6 from Jobin-Yvon (Longjumeau, France) with a SFM-3 stopped-flow module and a water circulating temperature control at 20 °C. The ellipticity at a selected wavelength was detected through a flow cell (FC15, 1.5 mm, 31 μL) for stopped-flow mixing or a cylindrical cell for manual mixing. Controls were recorded as described in Stopped-Flow Fluorescence Spectroscopy.

(a) Association/Folding. The kinetic measurements were obtained as described in Stopped-Flow Fluorescence Spectroscopy. The kinetic traces at 222 nm of 20 (0−100 s) and

[1] Abbreviations: CD, circular dichroism; NMR, nuclear magnetic resonance; Trx, thioredoxin; GuHCl, guanidine hydrochloride; UV, ultraviolet, SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis.

4 (100−1000 s) successive injections were recorded, averaged, and pasted.

(b) Refolding of the Denatured Trx. The kinetic measurements were obtained as described in Stopped-Flow Fluorescence Spectroscopy. The final concentration of Trx was 20 and 70 μM for the ellipticity measurements at 222 and 280 nm, respectively. The kinetic traces at 222 nm of 40 (0−2 s), 20 (2−20 s), and 10 (20−300 s) successive injections and the ones at 280 nm of 80 (0−20 s) and 20 (20−300 s) successive injections were recorded, averaged, and pasted.

(c) Reconstruction of Far-UV CD Spectra. The refolding was initiated by a manual 40-fold dilution of 0.8 mM previously denatured Trx, as described in Stopped-Flow Fluorescence Spectroscopy, with 2.34 mL of KP_i in a test tube under vigorous shaking. This solution was then rapidly transferred into a cylindrical cell of 2 mm path length. The recording was immediately triggered. The delay time between the beginning of mixing and the recording was 25 s. The kinetic traces of 4 (0−1800 s) successive additions at each selected wavelength (202−240 nm) were recorded, averaged, and fitted to a biexponential equation. Assuming a sequential process, the initial amplitude was calculated from the back-extrapolation to −25 s of the fitted slowest phase.

Data Analysis. Processing and fitting of the averaged kinetic traces were performed with the Figure P software from Biosoft (Cambridge, U.K.) using a multiexponential decay:

$$I = \sum_i A_i e^{-k_i t} + A_\infty \qquad (1)$$

where A_i is the amplitude of each phase, k_i its rate constant, and A_∞ is the plateau value. The observed rate constant of association, k_a, should increase linearly with the fragment, which is in excess, indicating that it corresponds to the association step.

$$N + C \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} NC^* \qquad k_a = k_{off} + k_{on}[C] \qquad (2)$$

1D-NMR Spectroscopy. In KP_i, 0.1 mM complex, 90% D_2O/10% H_2O, pH 5.7 (uncorrected) was prepared. A total of 1000 scans were acquired on a 500 MHz Varian Unity plus spectrometer using a S_2pul sequence, a spectral width of 8000 Hz, a 70° pulse angle, and 20 °C for data acquisition. The experiments were processed on a Silicon Graphics computer using Felix2.3 (Biosym Technologies, San Diego). The proton chemical shifts were calibrated on the water signal.

Sedimentation Diffusion Equilibrium. Solutions of N-fragment (A_{280} = 0.36), C-fragment (A_{230} = 0.38), and its equimolar mixture (A_{280} = 0.32) in KP_i were centrifuged at 20 °C in a Beckman XLA analytical ultracentrifuge at 35 000, 48 000, and 30 000 rpm, respectively, until equilibrium was reached. Each cell was scanned 10 times. The averaged scans were analyzed and fitted to a model for a single ideal species using the Data-Red Origin Single software from Beckman with a fixed baseline determined by high-speed meniscus clearing. The partial specific volumes, estimated from the aminoacid composition, were 0.717 mL/g (N-fragment) and 0.73 mL/g for the C-fragment and complex.

RESULTS

Association/Folding: Fluorescence and Far-UV CD. Stopped-flow measurements of the fluorescence (W28 and W31) were obtained using an excess of C-fragment. Analysis of a typical kinetic trace using 91 μM of one fragment and 2 μM of the other gave three exponential phases: $k_a = (12 \pm 2) \times 10^{-2}$ s^{-1}, $k_1 = (45 \pm 4) \times 10^{-3}$ s^{-1}, and $k_3 = (17 \pm 0.1) \times 10^{-4}$ s^{-1}. These type of measurements show a pseudo-first-order phase (see Figure 1A), a second phase which is masked by the former one at concentrations of C-fragment below 55 μM, and an apparently unimolecular slowest phase.

Manual measurements (see Figure 1, panels C and D, and Table 1) using an excess of either fragment are indeed consistent with an association step and a unimolecular slowest phase, whose precise value is difficult to determine and ranges between 0.0017 and 0.0033. Comparison of the measurements in the excess of either fragment indicate that an increase of the N-fragment concentration to 26 μM decreases the observed rate constant of association. The manual measurements show a decrease of 44% while the stopped-flow measurements indicate a greater decrease (66%). No decrease is, however, observed for the P76A variant. Analysis of the slowest phase of those measurements using an excess of the fluorescent N-fragment is difficult because (i) the quenching effect is much smaller than the total signal; (ii) this phase is masked by the photoeffect on the free N-fragment; (iii) the inner-filter effect introduces nonlinearities; (iv) the range of useful concentrations is limited, it should be high enough to distinguish the phases, but low enough to have observable relative changes.

If all the quenching were due to the formation of 1 μM complex, the relative quenching derived from the fits to the manual kinetic traces shown in Figure 1D (3.4% for the wild-type and 2.2% for the P76A mutant) would be expected to be 26 times smaller than that corresponding to Figure 1C (39% for wild-type, 22% for mutant). The nonlinearities might be expected to lower it even further. However, the opposite is observed, and neither the small drift due to the photoeffect nor the signal due to buffer alone can account for this discrepancy. This leads us to conclude that another process besides the reconstitution of the native structure is taking place. The stopped-flow measurements show even a higher relative quenching than the manual ones. It is plausible that this increase might be due to the type of measurement, which corresponds to the area under the fluorescence peak (stopped-flow) and the fluorescence at a single wavelength (manual).

Stopped-flow measurements of the change in ellipticity at 222 nm using high concentration of C-fragment (see Figure 1B) fit a three-exponential equation with rate constants similar to those observed by fluorescence: $k_a = (13 \pm 5) \times 10^{-2}$ s^{-1}, $k_1 = (30 \pm 2) \times 10^{-3}$ s^{-1}, and $k_2 = (15 \pm 3) \times 10^{-4}$ s^{-1}. The high relative amplitude of the bimolecular phase (82%) implies that most of the secondary structure is recovered in that step, although the slowest phase shows a negative relative amplitude (−38%). Together, stopped-flow and manual measurements, with exception of the ones in the excess of N-fragment, indicate an observed rate constant of association linearly dependent on the fragment in excess (Figure 1A) with apparent rate constants of association and

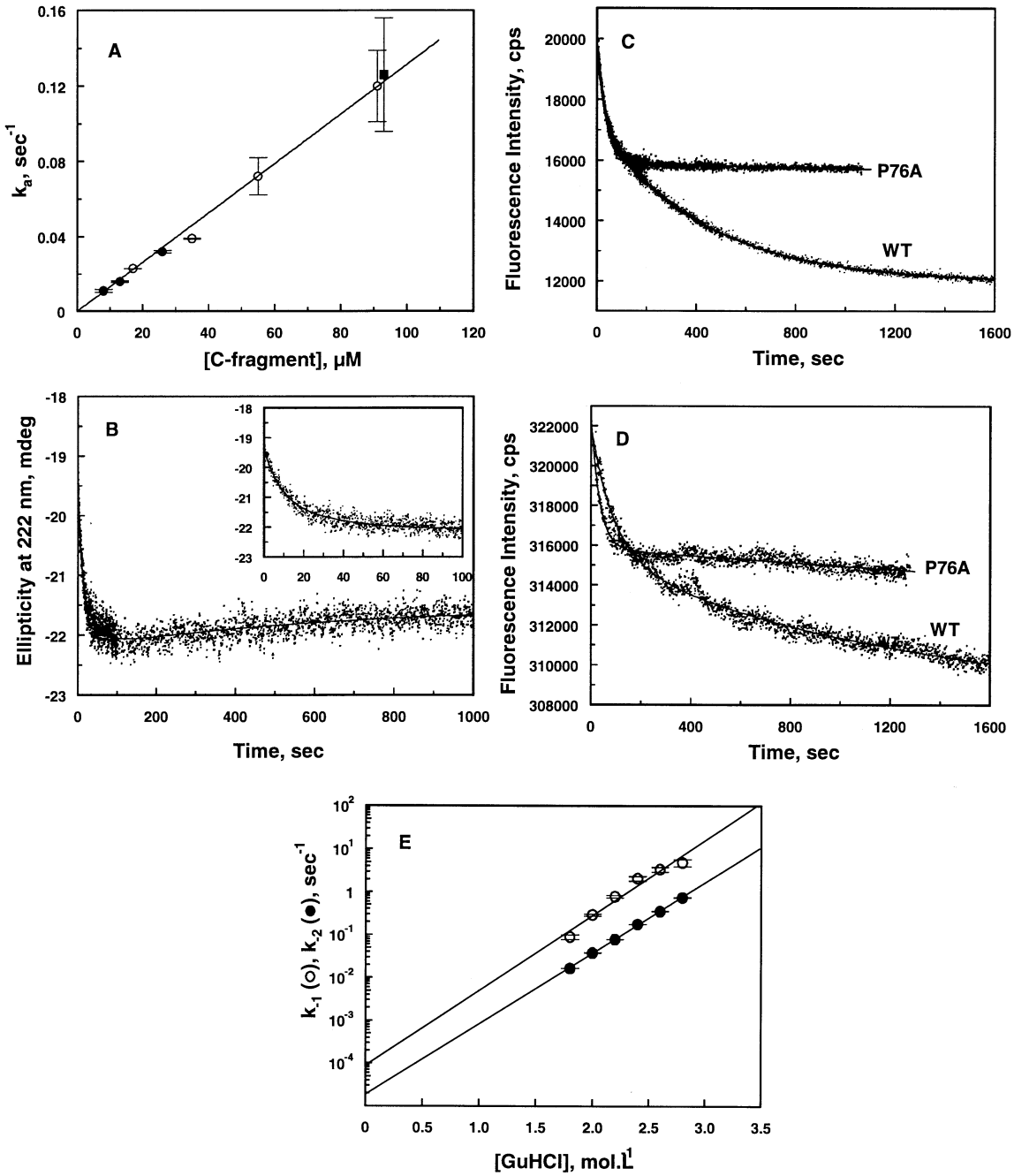


FIGURE 1: Solid lines show fits of the kinetic traces of association/folding to a multiexponential (A, B, and C). In panels C and D, WT means N-plus c-fragment and P76A means N-fragment plus P76A variant of the c-fragment. (A) Pseudo-first-order rate constants of association (k_c) from stopped-flow far-UV CD (filled square), stopped/flow fluorescence (empty circle), and manual fluorescence (filled circle) versus the concentration of C-fragment. Each error bar represents 1 standard deviation. (B) Stopped-flow far-UV CD kinetic trace using N-fragment (7.5 μM) and C-fragment (93 μM). The inset shows the first 100 s. (C) Overlay of kinetic traces from manual fluorescence using N-fragment (1 μM) and C-fragment (26 μM), or P76A (26 μM). (D) Overlay of kinetic traces from manual fluorescence using N-fragment (26 μM) and C-fragment (1 μM), or P76A (1 μM). (E) Rate constants of denaturation of the complex for the first phase, k_{-1} (empty circles), and second phase, k_{-2} (filled circles), versus the concentration of GuHCl. Each error bar represents 1 standard deviation.

dissociation, $k_{on} = 1{,}330 \pm 54$ M^{-1} s^{-1} and $k_{off} = 10^{-9}$ M^{-1} s^{-1}, respectively (see eq 2). The two slowest phases are best fit with $k_1 = 0.037 \pm 0.007$ s^{-1} and $k_2 = 0.0020 \pm 0.0005$ s^{-1}. Stopped-flow kinetic measurements using an excess of the P76A variant (100 μM) fit also three phases [$k_a = (10 \pm 0.4) \times 10^{-2}$ s^{-1}, $k_1 = (32 \pm 0.9) \times 10^{-3}$ s^{-1}, and $k_2 = (24 \pm 2) \times 10^{-4}$ s^{-1}], but the relative amplitude of the slowest phase decreases from 52 to 29%. At lower concentrations of the variant (see Table 1 and Figure 1C), this phase is replaced by a yet nonunderstood slower phase with a small relative negative amplitude. The same behavior is observed when the other fragment is in excess (see Figure 1D) but the observed rate constant of associations remain unchanged.

Dissociation/Unfolding: Fluorescence. Stopped-flow kinetic measurements of the denaturation of the complex were carried out at a concentration 100 times above the K_d and increasing concentrations of GuHCl, based on the denaturation under equilibrium conditions (Georgescu et al., unpublished results). Fitting of those traces obtained using between 1.8 and 2.8 M GuHCl to a biphasic process reveals two phases where the faster one has smaller amplitude. Extrapolation of the rate constants of denaturation at zero GuHCl concentration (see Figure 1E) according to $-RT \ln k_i(\text{GuHCl}) = -RT \ln k_i(0) + m^{\#}_i[\text{GuHCl}]$ gives $k_{-1} = (9 \pm$

Table 1. Kinetic Parameters from the Manual Fluorescence Measurements of the Association/Folding Process

[N][a] (μM)	[C][a] (μM)	k_a [b] ($s^{-1} \times 10^{-3}$)	k_2 [c] ($s^{-1} \times 10^{-4}$)	k_3 [c] ($s^{-1} \times 10^{-5}$)	A_a [d] (%)	A_2 [d] (%)	A_3 [d] (%)
			Wild-Type C-Fragment				
1	26	32 ± 0.6	24 ± 0.1		31 ± 1	69 ± 0.2	
1	13	16 ± 0.4	19 ± 0.09		14 ± 0.1	86 ± 0.05	
1	8	11 ± 0.8	17 ± 0.1		7 ± 0.1	93 ± 0.06	
26[e]	1	18 ± 1	33 ± 2		49 ± 2	51 ± 3	
			P76A variant				
1	26	26 ± 0.2			100 ± 0.5		
1	13	19 ± 0.1		36 ± 1	108 ± 0.3		-8 ± 0.06
1	8	14 ± 0.8		39 ± 2	111 ± 0.4		-11 ± 0.08
26[e]	1	29 ± 0.2			100 ± 2		
13[e]	1	15 ± 0.4			100 ± 2		
8[e]	1	11 ± 0.3		36 ± 5	114 ± 1		-14 ± 0.3

[a] Final concentration of N- and C-fragment. [b] Observed rate constant of association. [c] Second phase rate constants. [d] Relative amplitudes. [e] The averaged kinetic traces were fitted including a constant drift to account for the photoeffect on the free N-fragment.

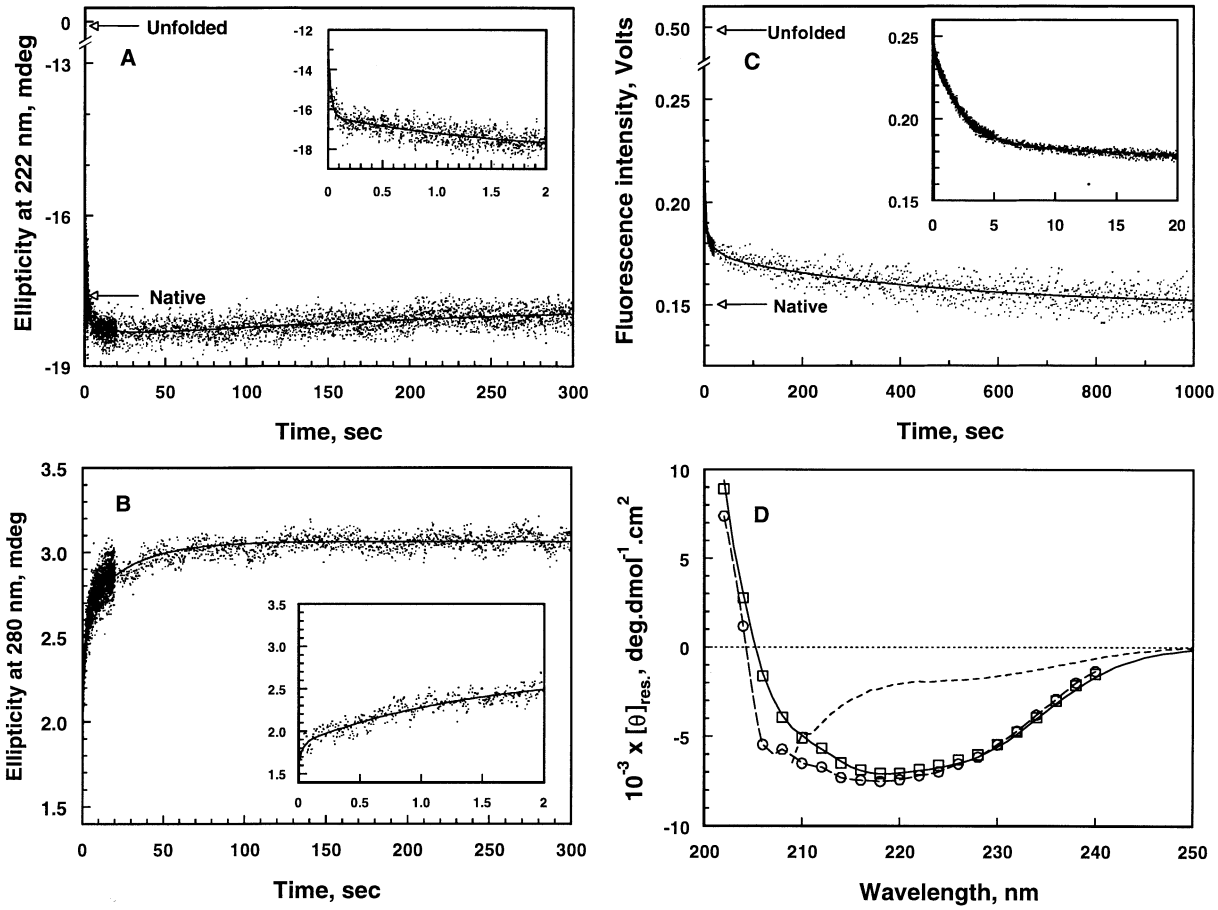


FIGURE 2: Refolding of denatured Trx. Solid lines show the fits of the stopped-flow kinetic traces to a multiexponential (A, B, and C). (A) Kinetic trace using far-UV CD. The ellipticity values at 222 nm for the unfolded and native Trx are indicated by arrows. The inset shows the first 2 s. (B) Kinetic trace using near-UV CD. The inset shows the first 2 s. (C) Kinetic trace using fluorescence. The inset shows the first 20 s. (D) Reconstructed far-UV CD spectra for the intermediate (empty circles) and Trx (empty squares). The spectrum of denatured Trx is shown in broken lines.

3) $\times 10^{-5}$ M^{-1} s^{-1}, $k_{-2} = (3.4 \pm 1.2) \times 10^{-5}$ M^{-1} s^{-1}, and $m^{\#}$ values of (4.0 ± 0.4) kcal mol^{-1} M^{-1} and (3.5 ± 0.1) kcal mol^{-1} M^{-1}, respectively.

Refolding of Denatured Trx. The refolding was monitored by far-UV CD (Figure 2A), near-UV CD (Figure 2B) and fluorescence (Figure 2C). Fitting of the kinetic traces for far-UV CD [>400 s^{-1}, $k_1 = (38.1 \pm 2.4)$ s^{-1}, $k_2 = (0.62 \pm 0.03)$ s^{-1}, $k_3 = (0.052 \pm 0.025)$ s^{-1}, and $k_4 = (29 \pm 19) \times 10^{-4}$ s^{-1}] and fluorescence [>400 s^{-1}, $k_1 = (33 \pm 2)$ s^{-1}, $k_2 = (0.429 \pm 0.003)$ s^{-1}, $k_3 = (0.040 \pm 0.003)$ s^{-1}, and $k_4 = (18 \pm 1) \times 10^{-4}$ s^{-1}] indicates the presence of a burst and four phases. The relative amplitudes of the two slowest far-UV CD phases show opposite signs (1.6% and -6%). The kinetic trace for near-UV CD [$k_1 = (33 \pm 2)$ s^{-1}, $k_2 = (0.70 \pm 0.03)$ s^{-1}, and $k_3 = (0.04 \pm 0.02)$ s^{-1}] fits a three-exponential equation without the burst and slowest phase. The three slowest phases agree with previous reports (Kelley & Stellwagen, 1984).

Refolding Intermediate: Far-UV CD. The same slowest rate constant was found by refolding denatured Trx at different wavelengths. Assuming a sequential pathway, the ellipticity values for Trx and its intermediate were derived and the spectrum reconstructed (see Figure 2D). Both spectra are similar, implying that this intermediate is rich in

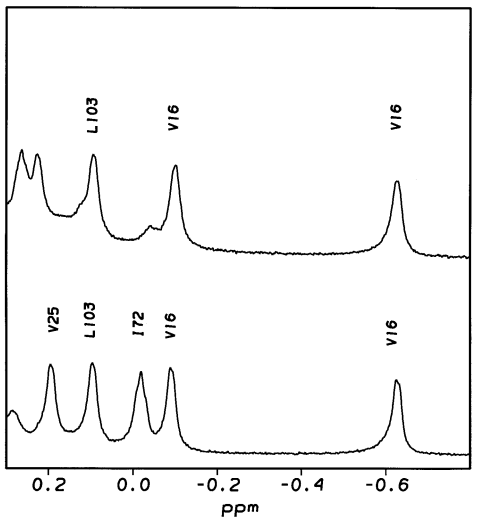


FIGURE 3: Upfield portion of the 1D-NMR spectrum of the complex between the N- and C-fragment (bottom trace) and its P76A variant (top trace).

secondary structure. Comparison between both spectra agrees with the overshoot in ellipticity seen in Figure 2A.

P76A Variant, NMR. Comparison of the 1D-NMR spectra of the complex and its P76A variant (see Figure 3) indicates upfield peaks typical of well-defined hydrophobic clusters. No variation in the packing of the $C^{\gamma}H$ of V16 against F81 and the $C^{\delta}H$ of L103 against Y49 is observed, but a significant downfield shift is found in the packing of the $C^{\delta}H$ of I72 against F27.

Monomers: Sedimentation. Fitting of the data from sedimentation-diffusion equilibrium to a single-species model indicates apparent molecular weights of 3810 (C-fragment), 7940 (N-fragment), and 10 690 (complex) which are comparable to those calculated for the corresponding monomeric species (3628, 8065, and 11 676, respectively).

DISCUSSION

Association/Folding vs Refolding. Fluorescence and far-UV CD were used to monitor the kinetics of the association/folding process under pseudo-first-order conditions. Comparison (see Table 1 and Figure 1) with the refolding of chemically denatured Trx (see Figure 2) indicates certain similarities. First, the two slowest phases have the same rate constants. In both cases, the fluorescence changes are accompanied by far-UV CD changes, suggesting that both processes involve similar structural changes around W28 and W31. Second, their relative amplitudes, although different, show the same tendencies (fluorescence quenching and recovery or loss of negative ellipticity at 222 nm, compare Figure 1B with 2A and Figure 1C with 2C). This agrees with the fact that the complex shows more intrinsic fluorescence (Reutimann et al., 1981), despite sharing an almost identical far-UV CD spectrum (Slabý & Holmgren, 1979). These differences might reflect local conformational changes in the protrusion that contains the W residues and the cleavage site (Tasayco & Chao, 1995). Despite these differences, the observed similarities lead us to believe that the fragments and denatured Trx share a common pathway that ends in the same overall structure.

Refolding Intermediate. Assuming a sequential pathway, a substantial recovery of native ellipticity at 222 and 280 nm occurs before the slowest phase, which suggests the formation of a compact intermediate rich in secondary (Figure 2D) and tertiary structure. The dependence of the slowest phase on the denaturation period, the formation of native-like-sized intermediates during the refolding of denatured Trx within a gel filtration column (Shalongo et al., 1987), and the disappearance of this phase during the refolding of the denatured P76A variant of Trx (Kelley & Richards, 1987) suggest that P76 isomerization occurs in this intermediate. This agrees with the following observations: (i) P76 isomerization might occur without perturbing the hydrophobic clusters due to its peripheral position in the β-sheet, and (ii) the upfield shift portion of the 1D-NMR spectrum (see Figure 3) of the P76A variant of the complex is typical of a compact structure with defined side chain packing.

Association/Folding Pathway. If we assume that the association/folding follows a sequential pathway, the problem becomes one of establishing the order of the events, in particular, whether the slow events occur before or after association. For instance, slow conformational changes, such as Pro isomerization, produce heterogeneities in the denatured state (Howry & Scheraga, 1996). If these slow changes occur in the fragments, they might produce some species that are not competent for association. Ultimately, these noncompetent species must become competent and form a complex since (a) an equimolar mixture at concentrations well above the K_d shows an almost identical far-UV CD spectrum to the one for the native uncleaved Trx, and (b) sedimentation analysis and molecular sieve chromatography of this mixture reflect a negligible amount of free fragment. If a slow phase is due to this conversion, an excess of the fragment with a high content of noncompetent population should result in (i) a decrease in the observed rate constant of association, since only the competent population is relevant, and (ii) a decrease in the relative amplitude of the slow phase. A slow phase occurring after association would not be affected. An analysis of the kinetic trace using 26 μM of the N-fragment under the assumption of pseudo-first-order conditions (see Table 1) indicates that an excess of this fragment decreases substantially the observed rate constant of association of the wild type C-fragment but not of its P76A variant. This selective effect suggests that indeed the N-fragment exists in both competent (N_{co}) and noncompetent (N_{nco}) monomeric forms. Assuming that all the trans/cis P76 isomers of the monomeric C-fragment present in a 10/1 ratio (Yu et al., unpublished results) are competent and that both of them associate with the same k_{on}, then the N_{co} amounts to 52% of the total. The possible presence of more than one unimolecular process with similar time scale (Mann et al., 1995), such as one occurring before association (possible the conversion of N_{nco} into N_{co}) and one occurring after, cannot be confirmed due to the above mentioned difficult analysis of the slowest phase.

The latter happens on the same time scale as the slowest refolding phase of the denatured uncleaved Trx, previously linked to the trans/cis P76 isomerization (Kelley & Richards, 1987). If it actually corresponds only to P76 isomerization, the P76A variant should eliminate it (Kelley & Richards, 1987). Indeed, kinetic studies of the association between the N-fragment (less or equal 26 μM) and the P76A variant show the elimination of this phase (see Figure 1, panels C and D, and Table 1) without changing the observed rate constant of association, regardless of which fragment is in

excess. These results imply that (i) the C-fragment and its variant have the same population of association competent species and (ii) the complete conversion of N_{nco} into N_{co} is not necessary for its productive association with the P76A variant. The fact that the C-fragment but not its variant distinguishes two major states of the N-fragment suggests that P76, which is located at the beginning of β-strand 4 in the native structure, might be related to zippering of the fragments through β-strands 2 and 4. The dramatic effect of the P76A variant on this phase and its little effect on the packing of the hydrophobic clusters of the complex (see Figure 3) strongly suggest that P76 isomerization occurs in a native-like complex with defined tertiary structure.

In summary, the native-like structural features of the complexes, the similarity of the slowest steps for the association/folding and refolding processes, the existence of noncompetent species of the N-fragment, and the kinetic effect of the P76A variation indicate that the reassembly of either the fragments or the uncleaved polypeptide chain into a native structure share common steps, which might involve conformational changes in the N-terminal region. Assuming that the C-fragment is completely competent, the isolated N-fragment appears to have at least 48% N_{nco}, which converts to N_{co} with a unimolecular rate constant k_{nco}. This specie binds to C^t and forms a native-like complex (NC^t) with tertiary structure, which then isomerizes to another native-like complex (NC^c) with a rate constant k^{tc} of the same order as k_{nco}.

Dissociation/Unfolding. Two phases are observed. Both are denaturant dependent, but just the slower accounts for more than 90% of the total change. For the uncleaved protein, only one denaturation phase has been reported (Kelley et al., 1986), consistent with the predominance of the cis isomer (Shalongo et al., 1987; Dyson et al., 1989). Both phases have similar $m^{\#}$ values (4.0 and 3.5 kcal/mol) which suggest a similar exposure of hydrophobic surfaces in the transition state (Matouschek & Fersht, 1993). These phases might be due to the denaturation of NC^c and NC^t, since both the peripheral position of P76 in the uncleaved protein and the effect of the P76A variation on the upfield shift peaks of the cleaved protein are consistent with similar hydrophobic clusters. However, one cannot rule out the possible contribution of another conformation since the NC^t is not detected by NMR even after using C-fragment with ^{13}C-labeled P76 (data not shown). The P76A variation does not affect the observed rate constant of association, but increases 12-fold the apparent dissociation equilibrium constant (Georgescu et al., unpublished results). This leads apparently to a larger rate constant of dissociation. Whether this corresponds to the observed 3-fold increase (9×10^{-5} s^{-1} versus 3.4×10^{-5} s^{-1}) of the denaturation phase with smallest amplitude is still unknown.

Kinetic Model. The accumulated structural and kinetic evidence on the formation and denaturation of the complex and its P76A variant provides clues about their mechanisms. Various kinetic models might explain the experimental results; we propose here a simple one, which neglects the middle phase observed at high concentrations of C-fragment. This model blends the association/folding and the dissociation/unfolding pathways in a reversible way and does not consider the possible isomers of the four trans Pro in the N-fragment. Initially, the rate constants (k_{nco}, k^t_{on}, k^c_{-1}, k^{tc})

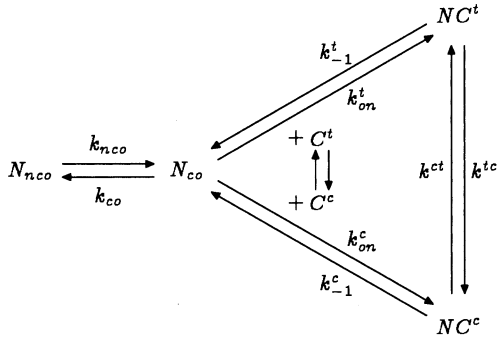
Scheme 1: Kinetic Model of Association/Folding

$$k_{nco} = 2 \times 10^{-3}\,s^{-1} \qquad k^{ct} = 2.1 \times 10^{-4}\,s^{-1}$$

$$k_{co} = 1.6 \times 10^{-3}\,s^{-1} \qquad k^{tc} = 4 \times 10^{-3}\,s^{-1}$$

$$k^t_{on} = 1.2 \times 10^{3}\,M^{-1}s^{-1} \qquad k^c_{on} = 2.5 \times 10^{3}\,M^{-1}s^{-1}$$

$$k^c_{-1} = 3.4 \times 10^{-5}\,s^{-1} \qquad k^t_{-1} = 3 \times 10^{-3}\,s^{-1}$$



in the model were assumed to be given directly by the measurements. The apparent rate constants of association (k^t_{on} and k^c_{on}) were assumed to be indistinguishable and the others were deduced to ensure consistency of the kinetic and steady state data with the model (see Supporting Information). To start with (see Scheme 1), there is a slow conversion of N_{nco} into N_{co} with a rate constant k_{nco} (2×10^{-3} s^{-1}). Both isomers C^t and C^c associate with N_{co} with the rate constants k^t_{on} and k^c_{on}, respectively. The same value, 1.3×10^3 M^{-1} s^{-1}, was used for both of them. The major isomer produces NC^t, which isomerizes with a rate constant k^{tc} (2×10^{-3} s^{-1}) to produce the major complex, NC^c. The latter is at least 19 times more abundant. This complex isomerizes back with a rate constant k^{ct} (1.1×10^{-4} s^{-1}). Both isomers dissociate/unfold with rate constants (at zero denaturant) k^t_{-1} (6.5×10^{-3} s^{-1}) and k^c_{-1} (3.4×10^{-5} s^{-1}) to produce N_{co}, which converts to N_{nco} with a rate constant k_{co} (1.8×10^{-3} s^{-1}).

We have run several simulations under the experimental conditions to relax the assumption of pseudo-first-order, which seems questionable if the N_{co} is just a fraction of the total N-fragment in excess. Using the program KINSIM (Barshop et al., 1983), one can obtain the concentration of each species as function of time for each experimental setup, based on the model. The molar signals of each species, however, are unknown and must be obtained by fitting to the observed kinetic traces. The simulation indicates that the assumed value for k^{tc} is too slow and for k^t_{-1} is too fast. A better fit is obtained with 4×10^{-3} s^{-1} and 3×10^{-3} s^{-1}, respectively (see Scheme 1). Consistency of the model forces k^{ct}, k^t_{on}, and k^c_{on} to change and seems to indicate that k^c_{on} is larger than k^t_{on}. If the difference between them is about 10-fold, they should be visible as two association phases in the excess of N-fragment, but our current manual measurements are not sensitive enough to assess the presence of a faster phase. To extract of N_{co}, we assume that the observed association is governed by C^t. This leads to a slightly higher value of 55% which determines the k_{co}.

The best fits, in which the molar signals are varied independently for each experiment, are indistinguishable from the observed kinetic traces in excess of either fragment

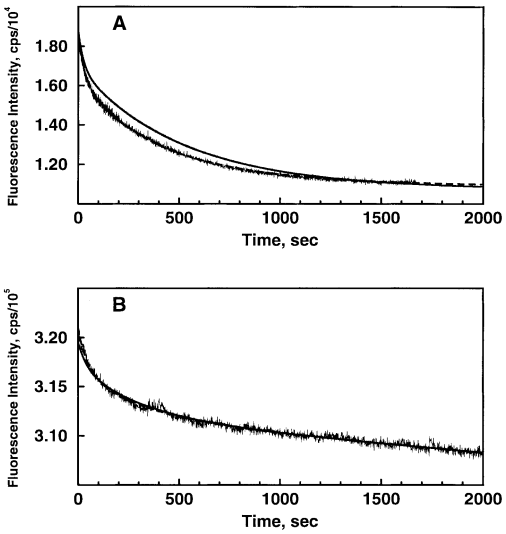


FIGURE 4: Three types of fits according to the model of Scheme 1 were superimposed on the kinetic traces of association/folding. Solid lines were used for the global fits. Short dashes were used for the best fits corresponding to 55% of N_{co}. (A) Overlay of global and best fits on the kinetic trace from manual fluorescence for N-fragment (1 μM) and C-fragment (26 μM). (B) Overlay of global and best fits on the kinetic trace from manual fluorescence for N-fragment (26 μM) and C-fragment (1 μM).

(see Figure 4). However, the relative molar signals obtained for each experiment are very different. For example, the ratio between the molar signals of NC^t and NC^c is 2.1 in excess of C-fragment and 0.95 in the inverse case. This is not an unexpected result, since the latter case seems to involve other processes besides the association/folding. A global fit, in which the relative molar signals for each species were assumed to be the same for all the experiments, was carried out at once for all the manual and stopped-flow biphasic kinetic measurements in excess of C-fragment. The other kinetic measurements were not considered to avoid the contribution of other processes. The values for the relative molar signals of N_{nco}, N_{co}, and NC^t were determined to be 3.3, 2.1, and 1.4 times the value for NC^c, respectively. In contrast, the same ratios for the best fit in Figure 4A (excess of C-fragment) are 1.26, 2.46, and 2.11, respectively. This discrepancy might be due to differences between the stopped-flow and manual measurements of the fluorescence as mentioned above. Figure 4 shows an overlay of the global and best fits on the manual kinetic traces using an excess of either fragment. The global fit is unexpectedly better for the manual kinetic trace using an excess of N-fragment which, presumably, involves other processes besides association/folding. We consider this an indication that a more complex kinetic model than the one presented here is necessary.

CONCLUSIONS

Our structural and kinetic studies indicate certain parallels between the association/folding and the refolding pathways: (i) the same overall final structure; (ii) the same two slowest rate constants coupling fluorescence and ellipticity changes; (iii) a slow step with an increase in the far-UV ellipticity in the slowest step and no apparent change in the overall secondary structure; and (iv) a trans/cis P76 isomerization apparently occurring in a compact intermediate with secondary and tertiary structure. This suggests a common pathway for both processes. These studies also reveal that (i) not all the N-fragment is competent to associate with the C-fragment; (ii) the noncompetent form seems to convert to competent on the same time scale as the P76 isomerization; (iii) the trans/cis P76 isomerization in the C-fragment is not required for recognition; (iv) the P76A variation does not alter the association competency of the C-fragment, but it permits its association with "noncompetent" forms of the N-fragment.

Studies of the refolding of denatured uncleaved proteins have the advantage of revealing in more detail the fast events, while studies of the reconstitution of a native-like structure using disordered protein fragments in nondenaturing conditions usually show a coupling between association and folding which might mask more than one event. However, the latter offers the advantage of recognizing slow events that occur in the disordered state, which points to regions of the uncleaved protein sequence that might produce misfolded states or noncompetent states for folding, information which is not readily revealed by native proteins.

ACKNOWLEDGMENT

We thank M. Amzel, C. Frieden, N. Kallenbach, and C. R. Matthews for their comments and F. E. Figueirido for helpful discussions. We also thank W.-F. Yu for the 1D-NMR spectra.

SUPPORTING INFORMATION AVAILABLE

Analysis of the kinetic model, three tables, analogous to Table 1, and results from sedimentation analysis and molecular sieve chromatography (8 pages). Ordering information is given on any current masthead page.

REFERENCES

Barshop, B. A., Wrenn, R. F., & Frieden, C. (1983) *Anal. Biochem. 130*, 134−145.
Chaffotte, A., Guillou, Y., Delepierre, M., Hinz, H.-J., & Goldberg, M. E. (1991) *Biochemistry 30*, 8067−8074.
Clark, A. C., Raso, S. W., Sinclair, J. F., Ziegler, M. M., Chaffotte, A. F., & Baldwin, T. O. (1997) *Biochemistry 36*, 1891−1899.
Dodge, R. W., & Scheraga, H. A. (1996) *Biochemistry 35*, 1548−1559.
Dyson, H. J., Holmgren, A., & Wright, P. E. (1989) *Biochemistry 28*, 7074−7087.
Freund, S. M., Wong, K. B., & Fersht, A. R. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 10600−10603.
Gittelman, M. S., & Matthews, C. R. (1990) *Biochemistry 29*, 7011−7020.
Howry, W. A., & Scheraga, H. A. (1996) *Biochemistry 35*, 11719−11733.
Katti, S. K., LeMaster, D. L., & Eklund, H. (1990) *J. Mol. Biol. 212*, 167−184.
Kelley, R. F., & Richards, F. M. (1987) *Biochemistry 26*, 6765−6774.
Kelley, R. F., & Stellwagen, E. (1984) *Biochemistry 23*, 5095−5102.
Kelley, R. K., Wilson, J., Bryant, C., & Stellwagen, E. (1986) *Biochemistry 25*, 728−732.
Kim, E. E., Varadajaran, R., Wyckoff, H. W., & Richards, F. M. (1992) *Biochemistry 31*, 12304−12314.
Kippen, A. D., & Fersht, A. R. (1995) *Biochemistry 34*, 1464−1468.
Labhardt, A. M. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 7674−7678.
Labhardt, A. M., Ridge, J. A., Lindquist, R. N., & Baldwin, R. L. (1983) *Biochemistry 22*, 321−327.

Langsetmo, K., Fuchs, J., & Woodward, C. (1989) *Biochemistry 28*, 3211−3220.
Lattman, E. E., Fiebig, K. M., & Dill, K. A. (1994) *Biochemistry 33*, 6158−6166.
Mann, C. J., & Matthews, C. R. (1993) *Biochemistry 32*, 5282−5290.
Mann, C. J., Shao, X., & Matthews, C. R. (1995) *Biochemistry 34*, 14573−14580.
Matouschek, A., & Fersht, A. R. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 7814−7818.
Milla, M. E., & Sauer, R. T. (1994) *Biochemistry 33*, 1125−1133.
Mulvey, R. S., Gualtieri, R. J., & Beychok, S. (1974) *Biochemistry 13*, 782.
Neira, J. L., Davis, B., Ladurner, A. G., Buckle, A. M., de Prat Gay, G., & Fersht, A. R. (1996) *Folding Des. 1*, 189−208.
de Prat Gay, G., & Fersht, A. (1994) *Biochemistry 33*, 7957−7963.
de Prat Gay, G., Ruiz-Sanz, J., & Fersht, A. (1994) *Biochemistry 33*, 7964−7970.
Reutimann, H., Luisi, P. L., & Holmgren, A. (1983) *Biopolymers 22*, 107.
Reutimann, H., Straub, B., & Luisi, P. L. (1981) *J. Biol. Chem. 256*, 6796−6803.
Robinson, C. R., & Sauer, R. T. (1996) *Biochemistry 35*, 13878−13884.
Sancho, J., & Fersht, A. R. (1992) *J. Mol. Biol. 224*, 741−747.
Schagger, H., & von Jagow, G. (1987) *Anal. Biochem. 166*, 368−379.
Shalongo, W., Ledger, R., Jagannadham, M. V., & Earle, S. (1987) *Biochemistry 26*, 3135−3141.
Shortle, D. R. (1996a) *FASEB J. 10*, 27−34.
Shortle, D. R. (1996b) *Curr. Opin. Struct. Biol. 6*, 24−30.
Slabý, I., & Holmgren, A. (1979) *Biochemistry 18*, 5585−5590.
Sosnick, T. R., Jackson, S., Wilk, R. R., Englander, S. W., & DeGrado, W. F. (1996) *Proteins: Struct., Funct., Genet. 34*, 427−432.
Taniuchi, H., & Anfinsen, C. B. (1971) *J. Biol. Chem. 246*, 2291−2297.
Tasayco, M. L., & Carey, J. (1992) *Science 255*, 594−597.
Tasayco, M. L., & Chao, K. (1995) *Proteins: Struct., Funct., Genet. 22*, 41−44.
Waltho, J., Feher, V., Merutka, G., Dyson, H. J., & Wright, P. E. (1993) *Biochemistry 32*, 6337−6347.
Wendt, H., Berger, C., Baici, A., Thomas, R. M., & Bosshard, H. R. (1995) *Biochemistry 34*, 4097−4107.
Zetina, C. R., & Goldberg, M. E. (1982) *J. Mol. Biol. 157*, 133−148.
Zitzewitz, J. A., Bilsel, O., Luo, J., Jones, B. E., & Matthews, C. R. (1995) *Biochemistry 34*, 12812−12819.

BI9708500