# Early Steps in Cytochrome *c* Folding Probed by Time-Resolved Circular Dichroism and Fluorescence Spectroscopy†

Gülnur A. Elöve,‡ Alain F. Chaffotte,§ Heinrich Roder,*,‡,∥ and Michel E. Goldberg§

*Institute for Cancer Research, Fox Chase Cancer Center, 7701 Burholme Avenue, Philadelphia, Pennsylvania 19111, Unité de Biochimie Cellulaire, CNRS URA 1129, Institut Pasteur, 28 rue du Dr. Roux, 75724 Paris Cedex 15, France, and The Johnson Research Foundation, University of Pennsylvania, Philadelphia, Pennsylvania 19104*

*Received March 31, 1992; Revised Manuscript Received May 11, 1992*

ABSTRACT: The kinetics of protein folding for horse ferricytochrome *c* was investigated by stopped-flow methods, using far-UV circular dichroism (CD), near-UV CD, and tryptophan fluorescence to probe the formation of secondary structure and tertiary interactions. In the far-UV region of the CD spectrum (222 nm), 44% of the total change associated with refolding occurs within the dead time of the stopped-flow experiment, indicating that a significant amount of helical secondary structure is formed in less than 4 ms. The remaining changes in the ellipticity at 222 nm occur in two kinetic phases with time constants of about 40 ms and 0.7 s, respectively. In contrast, there is no evidence for rapid changes in the ellipticity at 289 nm: an aromatic CD band, which is indicative of the formation of a tightly packed core, only begins to appear in a 400-ms step and is completed in a final 10-s phase. The fluorescence of a single tryptophan at position 59, which becomes quenched upon folding via nonradiative energy transfer to the heme group, provides complementary information on the condensation of the polypeptide chain during refolding. The fluorescence-detected stopped-flow folding kinetics of ferricytochrome *c* exhibits a 35% decrease in fluorescence during the dead time, suggesting that a substantial decrease in the average tryptophan–heme distance occurs on a submillisecond time scale. The subsequent fluorescence changes exhibit two prominent phases with time constants of about 20 and 300 ms, followed by a minor 5-s phase. Transient peptide CD spectra measured at different folding times (4 ms to 5 s) show no evidence for non-native elements of secondary structure at any stage of folding. Together with previous pulsed amide proton exchange data measured under identical folding conditions [Roder, Elöve, & Englander (1988) *Nature 335*, 700–704], the results suggest that during the early stages of folding (<4 ms), a partially condensed intermediate with a fluctuating core is formed that contains a significant amount of helical secondary structure but no stable hydrogen bonds. A second intermediate, populated on the 20–100-ms time scale, exhibits stable hydrogen-bonded structure in two helical segments near the chain termini and increased compactness, but the aromatic side chains are still in a dynamic or exposed environment. Other helical segments contribute to the helical character of the far-UV CD spectrum but lack sufficiently stable hydrogen bonding to provide protection against NH exchange.

Recent advances in the effort to unravel the mechanisms of protein folding have emphasized the importance of a detailed structural and kinetic description of folding intermediates (Kim & Baldwin, 1982, 1990; Ptitsyn, 1987; Kuwajima, 1989; Creighton, 1990; Matthews, 1991; Baldwin & Roder, 1991). This is especially true for the very early stages of folding, since some of the most crucial steps in directing the polypeptide chain toward its native conformation appear to occur very early in the folding process. Indeed, several structural features of the native conformation are already formed within the first 10–100 ms of the folding reaction (Udgaonkar & Baldwin, 1988; Roder et al., 1988; Bycroft et al., 1990; Briggs & Roder, 1992). Yet, a key question that remains unsolved concerns the relative importance of local versus long-range interactions in these early events. A better understanding of this question will require more detailed structural information on early folding intermediates.

A variety of conformational probes have been used to characterize transient folding intermediates. These include the following: far-UV circular dichroism to analyze secondary structure content (Pflumm et al., 1986; Labhardt, 1986; Kuwajima et al., 1987, 1991); the near-UV region of the CD[1] spectrum to detect the burial of aromatic groups in the anisotropic interior of the protein; pulsed amide proton exchange in conjunction with NMR to monitor the formation of stable hydrogen-bonded structure (Roder & Wüthrich, 1986; Udgaonkar & Baldwin, 1988; Roder et al., 1988; Bycroft et al., 1990; Miranker et al., 1991; Briggs & Roder, 1992); fluorescence properties such as energy transfer or spectral shifts to follow chain condensation and tertiary structure formation (e.g., Tsong, 1976; Blond & Goldberg, 1986; Garvey et al., 1989); immunoreactivity toward monoclonal antibodies

† This work was supported by grants from the National Institutes of Health (R01 GM35926 and CA06927 to H.R.), by an appropriation from the Commonwealth of Pennsylvania to the Institute for Cancer Research, and by funds from the Université Paris 7 and the Centre National de la Recherche Scientifique (URA 1129 to M.E.G.).

* To whom correspondence should be addressed.
‡ Fox Chase Cancer Center.
§ Institut Pasteur.
∥ University of Pennsylvania.

[1] Abbreviations: CD, circular dichroism; NMR, nuclear magnetic resonance; cyt *c*, cytochrome *c*; GuHCl, guanidine hydrochloride; ANS, 8-anilino-1-naphthalenesulfonate; BPTI, bovine pancreatic trypsin inhibitor.

0006-2960/92/0431-6876$03.00/0 

to detect the presence of nativelike structural patterns on the protein surface (Murry-Brelier & Goldberg, 1988; Blond-Elguindi & Goldberg, 1990). These methods provide complementary information, and their combined application is especially powerful. The far-UV region of the CD spectrum provides an estimate of secondary structure content and can indicate the presence of secondary structure elements even if they are unstable or non-native. However, the deconvolution of CD spectra in terms of $\alpha$-helices and $\beta$-structure is often problematic because of possible side-chain contributions and other complications (Sears & Beychok, 1979; Manning & Woody, 1989). Moreover, it only provides a global estimate of secondary structure content and gives no information on the location of $\alpha$-helices and $\beta$-strands within the structure. On the contrary, pulsed amide proton exchange followed by NMR analysis of the protected amide probes provides a precise description of the folding kinetics in terms of the formation of individual hydrogen bonds. However, only those amides that are protected against exchange in the native state can be observed, and one can only detect intermediates with stable hydrogen-bonded structure. This points to the complementarity of the pulsed amide proton exchange and kinetic circular dichroism approaches and prompted us to undertake a detailed comparison of the kinetics of secondary structure formation as monitored by the two methods.

The protein chosen for this study, horse cytochrome *c* (cyt *c*), is not only one of the best characterized in terms of its structure, both in the crystalline state (Bushnell et al., 1990) and in solution (Wand et al., 1989; Feng et al., 1989, 1990), but it also has served as a model protein for numerous equilibrium and kinetic folding studies (e.g., Babul & Stellwagen, 1972; Ikai & Tanford, 1973; Fisher et al., 1973; Knapp & Pace, 1974; Tsong, 1976; Ridge et al., 1981; Brems & Stellwagen, 1983). Moreover, the folding pathway of oxidized cyt *c* has been investigated in detail by pulsed amide proton exchange (Roder et al., 1988; Elöve & Roder, 1991). The overall picture emerging from these studies is that the folding reaction is complex and heterogeneous. About 50% of the molecules appear to follow a sequential pathway with a partially folded intermediate populated on the 10–100-ms time scale. Amide probes in two segments near the chain termini that correspond to two interacting $\alpha$-helices of the native structure were found to be protected in a common kinetic phase with a time constant of 20 ms, while amide probes on the intervening chain segment were protected only at a later stage of folding (Roder et al., 1988). These results indicate that two $\alpha$-helices near each end of the chain form and associate before stable hydrogen-bonded structure appears in other parts of the protein. The presence of two mutually interacting helices in the early intermediate was confirmed by the recent observation that N- and C-terminal amide sites not only are protected at the same rate but also exhibit the same degree of protection (Elöve & Roder, 1991). A subsequent folding step in the 100–500-ms time range involves changes in heme ligation (replacement of a non-native histidine ligand with the native methionine–iron ligand; Brems & Stellwagen, 1983; Elöve & Roder, 1991; Muthukrishnan & Nall, 1991) and coincides with the formation of the third major $\alpha$-helix of the native protein, the 60's helix (residues 61–70). In addition to the main folding pathway, there is evidence for a small fraction of very rapidly folding molecules that appear to form a nativelike structure within 4 ms and another minor species that takes several seconds to reach the native state. Fluorescence-detected stopped-flow experiments gave additional information on the overall conformation of the polypeptide chain (Roder et al., 1988), since the observed decrease in tryptophan fluorescence during folding is largely due to energy transfer between a single tryptophan, Trp 59, and the covalently bound heme group (Tsong, 1974, 1976; Jeng & Englander, 1991).

Kuwajima et al. (1987) previously applied stopped-flow circular dichroism methods to investigate the kinetics of secondary structure formation in horse cyt *c*. In refolding experiments performed at 25 °C, they found that over 80% of the change in the CD signal at 222.5 nm occurred within the 18-ms dead time of their mixing apparatus, whereas the heme-related CD signal at 420 nm showed no such rapid changes, suggesting that an intermediate with nearly native levels of helical secondary structure was formed in less than 18 ms. This conclusion is in apparent contradiction with the pulsed NH exchange results mentioned above (Roder et al., 1988). Indeed, on the basis of the time course of protection for helical amide probes, one would expect only about 30% of the native secondary structure to be formed within the first 20 ms of refolding. However, it is difficult to compare these studies, since they were carried out under different refolding conditions (10 °C, pH 6.2, and a final GuHCl concentration of 0.7 M for pulsed NH exchange compared to 25 °C, pH 6.8–7, and 0.4 M GuHCl for the kinetic CD studies).

In order to address this apparent discrepancy, we measured the kinetics of folding for oxidized cyt *c* by time-resolved CD spectroscopy under the conditions used in previous pulsed NH exchange studies (Roder et al., 1988), taking advantage of the short dead time (4 ms) and high sensitivity of our stopped-flow CD instrument. This permits a more direct comparison with folding studies based on other conformational probes, including fluorescence and protection against amide proton exchange.

## MATERIALS AND METHODS

*Materials and Sample Preparation.* Horse heart ferricytochrome *c* (type VI from Sigma Chemical Co.) was used without further purification. GuHCl was ultrapure grade from ICN-Schwarz/Mann. All other chemicals were reagent grade. Protein concentrations were determined spectrophotometrically, using an extinction coefficient of $1.06 \times 10^5\ M^{-1}\ cm^{-1}$ (Babul & Stellwagen, 1972). GuHCl concentrations were based on refractive index measurements (Pace, 1986) performed on a Reichert-Jung Mark II Abbe refractometer (Leica, Inc., New York). The pH was adjusted by addition of HCl or NaOH and measured before and after refolding experiments. Unfolded protein solutions were equilibrated for a minimum of 1 h prior to measurements.

*CD Instrument.* Equilibrium and time-resolved CD experiments were performed on a CD6 spectrodichrograph (Jobin-Yvon Instruments SA, Longjumeau, France) equipped with a stopped-flow accessory. This accessory consists of an SFM-3 stopped-flow module (Bio/Logic, Claix, France) interfaced with the CD6, using a kit provided by Jobin-Yvon. The observation cell used was a T50/15 transmission cell with a 5-mm optical path. The stopped-flow module and observation cell were thermostated at 10 ± 0.1 °C by means of a temperature-controlled water bath and a high-flow circulating pump.

*Equilibrium Measurements.* Equilibrium CD spectra were recorded at 10 °C in the far-UV (210–260 nm), near-UV (250–350 nm), and visible (350–500 nm) regions, both for unfolded (in 4.3 M GuHCl) and for refolded (in 0.73 M GuHCl and 100 mM acetate buffer) samples of oxidized cyt *c* (25 $\mu$M) at pH 6.4. The integration time was 2 s, and the

wavelength interval was 1 nm. At least three accumulations were averaged, and the spectrum of the corresponding solvent was subtracted.

*Stopped-Flow CD Measurements.* For all stopped-flow folding experiments, unfolded oxidized cyt *c* in 4.3 M GuHCl, pH 6.3, was diluted at 10 °C with 0.1 M acetate buffer, pH 6.3, to initiate refolding at a final GuHCl concentration of 0.7 M. The mixing protocol for CD-detected stopped-flow experiments varied with the wavelength of interest. In order to minimize mixing artifacts, a small volume of GuHCl-unfolded cyt *c* solution was loaded into the 5-mL syringe of the mixing module and mixed with a large volume of refolding buffer (containing some GuHCl), injected via two 18-mL syringes. For measurements at 222 nm, 10 μL of a 0.64 mM cyt *c* solution in 4.3 M GuHCl was mixed with 790 μL of 100 mM acetate buffer containing 0.66 M GuHCl, to give a final GuHCl concentration of 0.7 M and a final protein concentration of 8 μM. For measurements at 289 nm, 28 μL of a 2 mM cyt *c* solution in 4.3 M GuHCl was mixed with 812 μL of 0.1 M acetate containing 0.58 M GuHCl, yielding a final solution of 65 μM cyt *c* in 0.7 M GuHCl.

Kinetic traces (typically 1000 points) covering two or three different time bases ranging from 1 to 60 s were recorded at each wavelength, using sampling times of 1–60 ms and appropriate filter constants. The dead time was ~4 ms for 1-s traces and ~6 ms for longer traces (5–60 s), which were recorded at a lower flow rate. For each time base, 40–120 transients of 1000 data points were averaged.

Since each syringe of the Bio-Logic stopped-flow module can be controlled individually, it was possible to record all relevant baseline readings along with each refolding experiment, as previously described (Chaffotte et al., 1992). Each stopped-flow trace was preceded by successive injections of refolding buffer (0.1 M acetate, 0.66 M GuHCl), denaturant solution (4.3 M GuHCl), and unfolded cyt *c* (8 μM at 222 nm and 65 μM at 289 nm) in 4.3 M GuHCl, taking a 1-s ellipticity reading for each phase. The CD signal for the refolding buffer was subtracted from all other data to remove any instrumental offsets. The ellipticity of the fully unfolded protein, $\theta_u$, was determined by subtracting the denaturant background from the reading for cyt *c* in 4.3 M GuHCl. Since the ellipticity of unfolded cyt *c* was found to vary little with denaturant concentration (see Results), we used this value as an estimate for the ellipticity of the still unfolded protein under refolding conditions. The ellipticity for the refolded protein, $\theta_f$, was determined from the baseline of the kinetics at long times.

*Continuous-Flow CD Measurements.* Continuous-flow experiments were performed by using the following mixing protocol. The CD signal was first recorded after injecting 800 μL of refolding buffer to obtain a buffer reference value. During a subsequent 500-ms continuous-flow phase, 50 μL of unfolded cyt *c* (in 4.3 M GuHCl) was mixed with 3.95 mL of refolding buffer (containing 0.66 M GuHCl), and the ellipticity was determined by averaging 100 data points recorded at 5-ms intervals (5-ms filter constant). The flow rate was 8 mL/s, resulting in an effective folding time of 4 ms for the protein observed during the injection phase (i.e., delay between the mixer and the center of the observation cell).

*Fluorescence-Detected Kinetics.* Fluorescence-detected folding experiments were performed under very similar conditions on a PQ/SF-53 stopped-flow instrument (Hi-Tech, Salisbury, England) equipped with a Berger-type mixing chamber and a 2 × 2 × 10 mm flow cell. A 75-W xenon lamp (On-Line Instrument Systems, Inc., Jefferson, GA) and a monochromator (Hi-Tech) were used for excitation at 280 nm (5-nm bandwidth) along the 10-mm axis of the flow cell. The fluorescence emission was measured in the 2-mm direction, using high-pass glass filters with a 320- or 360-nm cutoff (filters WG-320 and WG-360 provided by Hi-Tech). Refolding was initiated at 10 °C by 6-fold dilution of oxidized cyt *c* in 4.2 M GuHCl with 0.1 M acetate buffer at pH 6.2, resulting in a final GuHCl concentration of 0.70 M and a final protein concentration of 35 μM. The fluorescence changes associated with refolding were recorded at a sampling time of 64 μs, using a personal computer equipped with a DT-2801/A (Data Translation, Inc., Marlboro, MA) data acquisition board and a program written in ASYST language by S. D. Luck and H. Roder.

*Kinetic Analysis.* Kinetic curves were constructed by combining the CD traces collected at different time bases (1000 points each), after correction for minor baseline offsets. The total number of points in the combined traces was reduced to about 170, using the following logarithmic averaging scheme. After the first 16 points, consecutive sampling points were averaged, increasing the number of averaged points toward longer times. Kinetic parameters were obtained by nonlinear least-squares fitting of two or three exponential phases, using a program based on the ASYST software package. The amplitudes for the individual phases were normalized with respect to the total change in ellipticity upon refolding, $\theta_u - \theta_f$, measured in the control experiments described above.

## RESULTS

*Folding–Unfolding Equilibrium.* The reversible GuHCl-induced unfolding transition monitored by tryptophan fluorescence for oxidized horse cyt *c* at 25 °C, pH 7, occurs between about 2 and 3.5 M GuHCl with a midpoint concentration, $C_m$, of 2.7 M (Tsong, 1974). Preliminary equilibrium CD measurements at 222 nm indicate a slightly higher $C_m$ of 3 M at 10 °C and pH 6.3. These experiments also showed that the ellipticity at 222 nm is virtually independent of the GuHCl concentration in the baseline region above 4 M. Thus, the protein is fully unfolded at 4.3 M GuHCl, the initial GuHCl concentration chosen for the present studies, and the final concentration, 0.7 M, lies well below the unfolding transition.

*CD-Detected Folding Kinetics.* Representative stopped-flow CD experiments monitored in the peptide (222-nm) and aromatic (289-nm) regions of the CD spectrum are shown in Figure 1. Panel A shows raw data for a 1-s trace of the kinetics at 222 nm recorded with a 1-ms sampling time and a dead time of 4 ms. The data represent the averaged signal from 120 recordings. Two additional experiments were performed under the same conditions but using sampling times of 5 and 25 ms to cover the time range up to 25 s (cf. Figure 1C). Folding experiments monitored at 289 nm were measured on a 5-s time scale (Figure 1B) and a 60-s time scale (not shown). For this data set, 38 traces of 1000 points were averaged.

The kinetic traces recorded on different time scales were combined and normalized with respect to the total ellipticity change, $\theta_f - \theta_u$, where $\theta_u$ represents the ellipticity for the unfolded protein (see Materials and Methods) and $\theta_f$ represents the native value measured at long times. The total number of sample points was reduced by logarithmic averaging of the combined traces. The normalized and averaged data at both wavelengths are presented in Figure 1C, using a logarithmic time scale in order to capture the complete kinetics over the time range from 4 ms to 60 s in a single graph. The solid lines represent nonlinear least-squares fits with two exponential

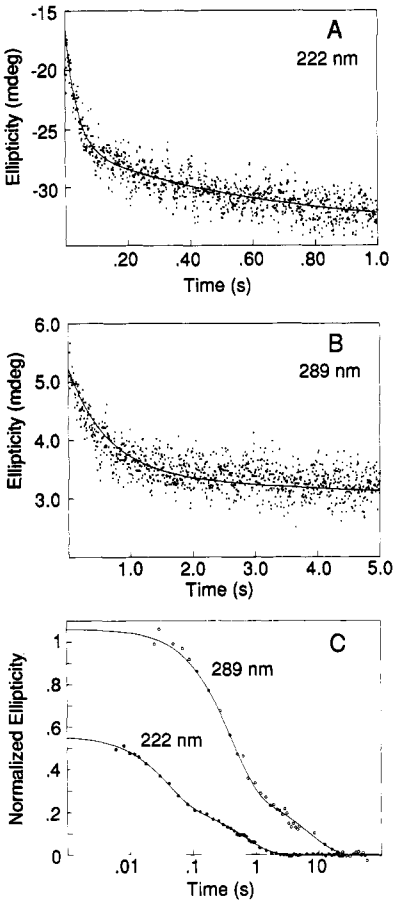


FIGURE 1: CD-detected stopped-flow folding kinetics for oxidized cyt *c* at pH 6.3, 10 °C, initiated by a GuHCl concentration jump from 4.3 to 0.7 M GuHCl. The time course of refolding was monitored at 222 nm (panel A) and 289 nm (panel B). The folding kinetics at both wavelengths is compared in panel C, which shows plots of the normalized ellipticity, relative to the fully unfolded state (fraction unfolded), on a logarithmic time scale. The data in panels A and B were combined with similar traces recorded over a longer time scale (up to 25 s to 222 nm and up to 60 s at 289 nm) and subjected to logarithmic averaging (see Materials and Methods). The curves represent double-exponential fits determined by nonlinear least-squares analysis. The resulting kinetic parameters are summarized in Table I.

phases (segments of the same fits are also shown on a linear time scale in panels A and B). The resulting rates and relative amplitudes are summarized in Table I.

The double-exponential fit to the 222-nm data (Figure 1C) extrapolates to 0.56 at $t = 0$, indicating that 44% of the total folding-related change in the peptide CD signal occurs within the mixing dead time of about 4 ms. Following Kuwajima et al. (1991), we will refer to this rapid folding process as the "burst phase", and for convenience, the subsequent folding events will be grouped into "fast phases" (4–100 ms), "intermediate phases" (100 ms–1 s), and "slow phases" (>1 s). The initial part of the kinetics observed at 222 nm corresponds to an exponential phase with a rate $k_1 = 26\ s^{-1}$ and a relative amplitude $A_1 = 0.33$ (Table I). This fast folding phase is followed by an intermediate phase with a rate $k_2 = 1.5\ s^{-1}$ and relative amplitude $A_2 = 0.23$. At the end of the intermediate phase, the peptide CD signal has essentially reached the native baseline value, in contrast to other probes (see below) which display additional slower phases.

The changes during refolding observed in the aromatic region of the CD spectrum differ markedly from those observed in the peptide region (Figure 1C). At early times of refolding, the ellipticity at 289 nm is in close agreement with the value measured for the fully unfolded protein and remains at this level for about 100 ms. Thus, the burst phase and the fast phase that dominate the kinetics at 222 nm are not detected at 289 nm. The major change (76% of the total amplitude) in the aromatic CD band occurs on the time scale of the intermediate phase ($k_2 = 2.5\ s^{-1}$), followed by a slower phase ($k_3 = 0.1\ s^{-1}$) for the remaining 29% of the total ellipticity change.

*Wavelength Dependence.* Stopped-flow folding experiments were performed at various wavelengths in the far-UV region of the CD spectrum in order to determine transient CD spectra at different stages of the cyt *c* folding reaction. The wavelength dependence of the CD signal at early (4 ms), intermediate (100 ms), and late (5 s) stages of refolding is compared in Figure 2 with equilibrium CD spectra for fully unfolded cyt *c* in 4.3 M GuHCl (trace a) and native cyt *c* in 0.7 M GuHCl (trace d). Due to the high absorbance background of the denaturant and acetate buffer, the signal-to-noise ratio decreased rapidly at low wavelengths, and the number of averaged kinetic traces was insufficient to consider data below 218 nm. The data points at 100 ms and 5 s were obtained from multiexponential fits to the kinetic traces, and the transient spectrum at an effective folding time of 4 ms was measured in continuous-flow experiments at various wavelengths (see Materials and Methods).

The transient spectrum measured at 5 s agrees well with the independently determined equilibrium spectrum for native cyt *c* (in the same buffer), confirming the earlier conclusion that, in the far-UV region of the CD spectrum, the protein appears fully folded after the intermediate phase, at a stage where the protein has not reached the native state according to other parameters, including the ellipticity at 289 nm (Figure 1). The dashed lines in Figure 2 show that the wavelength dependence at shorter times can be interpreted in terms of linear combinations of the equilibrium spectra for the folded (N) and the unfolded (U) forms with 40% N/60% U at 4 ms and 72% N/28% U at 100 ms. This suggests that there is no measurable contribution from intermediates with non-native secondary structure.

*Fluorescence-Detected Folding Kinetics.* The folding kinetics for oxidized cyt *c* monitored by tryptophan fluorescence is compared in Figure 3 with the far-UV CD measurements and the results from previous pulse labeling experiments (Roder et al., 1988). The fluorescence emission above 310 nm, normalized relative to the unfolded state under refolding conditions, is plotted on a logarithmic time scale. The fluorescence signal of the unfolded protein, $F_u$(4.2 M), was measured by mixing unfolded cyt *c* (in 4.2 M GuHCl) with a 5-fold excess of a 4.2 M GuHCl solution, instead of the refolding buffer. However, the fluorescence for fully unfolded cyt *c* varies with GuHCl concentration, increasing by about 25% from 4 to 6 M GuHCl (Tsong, 1974), which has to be taken into account in estimating the fluorescence of the unfolding protein under refolding conditions, $F_u$(0.7 M). The relative fluorescence intensity was determined by normalizing the raw data with $F_u$(0.7 M) – $F_n$(0.7 M), where $F_n$(0.7 M) represents the signal of the refolded protein. The fact that a very similar kinetic trace was obtained in a parallel experiment in which the fluorescence emission was measured above 350 nm indicates that the emission maximum remains near the value of the fully unfolded state (356 nm) throughout the folding reaction.

The fluorescence changes associated with refolding are interpreted in terms of four kinetic phases (Table I). About 35% of the total fluorescence change occurs within the mixing

Table I: Kinetic Parameters for Refolding of Oxidized Cytochrome *c*[a]

| probe | $k_0$ ($s^{-1}$) | $A_0$ | $k_1$ ($s^{-1}$) | $A_1$ | $k_2$ ($s^{-1}$) | $A_2$ | $k_3$ ($s^{-1}$) | $A_3$ |
|---|---|---|---|---|---|---|---|---|
| far-UV CD | >250 | 0.44 | 26 ± 5[b] | 0.33 ± 0.03 | 1.5 ± 0.5 | 0.23 ± 0.02 | | |
| near-UV CD | | | | | 2.5 ± 0.8 | 0.76 ± 0.07 | 0.10 ± 0.09 | 0.29 ± 0.04 |
| fluorescence | >250 | 0.35 | 56 ± 6 | 0.31 ± 0.01 | 3.2 ± 0.3 | 0.29 ± 0.01 | 0.18 ± 0.06 | 0.05 ± 0.01 |
| N- and C-helices[c] | >250 | 0.17 | 60 | 0.43 | 5.0 | 0.25 | 0.10 | 0.15 |
| other NH probes[c] | >250 | 0.14 | | | 5.0 | 0.37 | 0.10 | 0.49 |

[a] From stopped-flow or quenched-flow folding experiments at 10 °C in 0.7–0.72 M GuHCl and 0.1 M acetate, pH 6.2. Amplitudes are normalized relative to the total signal change between fully unfolded and folded states. $A_0$ is the amplitude of fast processes occurring in the dead time, based on extrapolation of the signal to $t = 0$. [b] Standard error (95% confidence limit) from nonlinear least-squares analysis. [c] From pulsed NH exchange experiments (Roder et al., 1988).

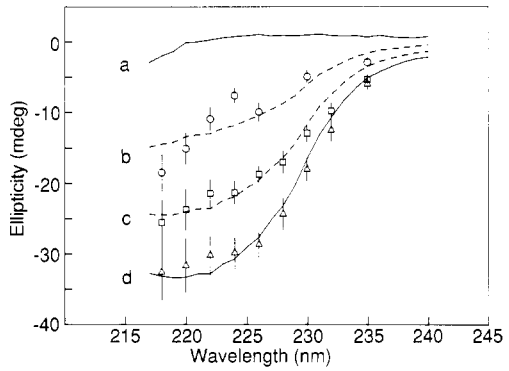


FIGURE 2: Comparison of transient peptide CD spectra with equilibrium spectra of native and unfolded oxidized cyt *c* (8 μM). The data points show the wavelength dependence at 100 ms (□) and 5 s (△), determined by constrained nonlinear least-squares fitting of the stopped-flow kinetics (fixed rates). Continuous-flow data at an effective folding time of about 4 ms are also shown (○). Equilibrium CD spectra of unfolded cyt *c* in 4.3 M GuHCl (a) and native cyt *c* in 0.72 M GuHCl (d) are drawn as solid curves. The dashed curves indicate linear combinations of the spectra of the folded (F) and the unfolded (U) forms as follows: 0.40 F + 0.60 U at 4 ms (b); 0.72 F + 0.28 U at 100 ms (c).

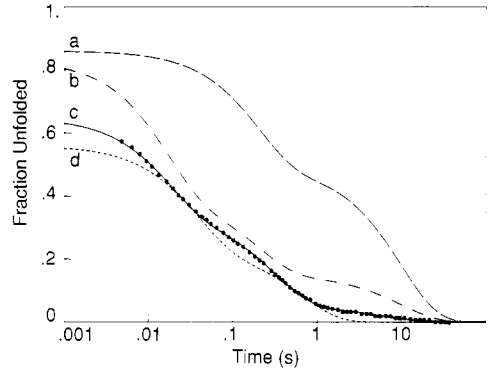


FIGURE 3: Fluorescence-detected folding kinetics of oxidized cyt *c* (●) following a GuHCl concentration jump from 4.2 to 0.7 M at pH 6.2, 10 °C. The fluorescence emission was measured above 310 nm and normalized relative to that of unfolded protein under the same conditions. The solid curve (c) represents sums of three exponential phases determined by least-squares analysis (Table I). For comparison, the folding kinetics monitored by other conformational probes is also shown on the same logarithmic time scale. The top two curves represent the time course of protection for two distinct classes of amide protons observed by pulsed NH exchange (Roder et al., 1988); amide probes in the 60's and the 70's helices and those involved in tertiary H-bonds (a) and amide probes in the N- and C-terminal helices (b). Curve d represents the folding kinetics monitored by far-UV CD (cf. Figure 1).

dead time (burst phase). An additional 30% decrease occurs in a fast phase with a time constant of 18 ms, followed by another 30% decrease on the time scale of the intermediate phase (~300 ms) and a minor slow phase (5 s). These parameters are similar to those reported by Roder et al. (1988), except for $A_0$, which was not determined previously. Some variation in kinetic parameters is expected due to their strong dependence on the conditions of refolding (especially pH and denaturant concentration).

The fluorescence intensity of unfolded cyt *c* in 4.2 M GuHCl is about 50% of that for free tryptophan under the same conditions and essentially 0 for the folded protein (Tsong, 1974). The fluorescence decay upon folding has been attributed to Förster energy transfer between Trp 59 and the heme group (Tsong, 1974, 1975), which results in complete quenching of the fluorescence in the folded state where the Trp 59 ring is in close proximity with the heme group (9.5-Å center-to-center distance). Tsong (1975) described an acid-induced increase in the fluorescence of Trp 59 from 55% to over 90% relative to free tryptophan for urea-unfolded cyt *c* with a parallel increase in reduced viscosity from 15.5 to 23 mL/g. Similarly, Jeng and Englander (1991) observed a salt-dependent decrease in fluorescence intensity at a constant emission maximum (356 nm) for acid-denatured cyt *c* with a simultaneous decrease in molecular volume as measured by viscometry. Our observation that the folding kinetics is independent of the emission wavelength range monitored indicates that fluorescence quenching is dominated by tryptophan–heme energy transfer at all stages of refolding. Due to the inverse sixth power dependence of the transfer efficiency on donor–acceptor distance (Förster, 1974), the fluorescence intensity thus represents a sensitive measure of the average distance between Trp 59 and the heme.

## DISCUSSION

The observation of multiple kinetic phases when a particular conformational probe is monitored is not necessarily indicative of the presence of intermediates on the folding pathway: complex folding kinetics can also result from a heterogeneous population of unfolded molecules caused, for example, by cis–trans isomerization of proline peptide bonds and, in the case of cyt *c*, by alternative heme ligands (Brems & Stellwagen, 1983; Elöve & Roder, 1991). However, evidence for structural folding intermediates can be obtained by comparing the kinetics observed with different probes. In the present case, the large differences in amplitudes and rates observed in the far-UV and near-UV regions of the CD spectrum (Figure 1) provide clear evidence for the transient accumulation of partially folded intermediates. Further information on folding intermediates can be obtained by comparing the results of this study with previous pulsed NH exchange results (Figure 3).

*The Burst Phase (<4 ms).* About 45% of the overall change in the far-UV CD signal occurs within the instrumental dead time, indicating that a significant amount of secondary structure is formed within the first 4 ms of refolding. The transient CD spectrum measured after the burst phase (4-ms trace in Figure 2) is consistent with partially helical structure, although, given the uncertainty of these measurements, minor

contributions from non-native secondary structure cannot be excluded (see below). The 35% decrease in fluorescence intensity relative to the unfolded state observed during the dead time (Figure 3) corresponds to a significant reduction in the average tryptophan–heme distance and is consistent with a partial collapse of the polypeptide chain at the very early stages of folding. On the other hand, the aromatic CD signal does not exhibit a burst phase and remains at the level of the unfolded state out to about 100 ms. At this stage, the aromatic side chains are apparently still in a largely isotropic environment, due to either solvent exposure or motional averaging.

Previous pulsed NH exchange studies (Roder et al., 1988) showed that most of the observed amide protons acquire 10–20% protection during the 3-ms dead time of the quenched-flow experiment. The small amplitude of this burst phase, compared to the far-UV CD data (Figure 3), indicates that, at 4 ms, the secondary structure detected by CD is insufficiently stable to protect amide protons against exchange. This is not suprising, since only amide protons involved in stable H-bonded structure are protected from exchange, while even marginally stable helices can result in a CD spectrum with partially helical character.

These observations indicate that the folding intermediate formed during the burst phase is a partially condensed state with elements of secondary structure, consisting of marginally stable, fluctuating helices but without stable H-bonds or a tightly packed core. These are some of the attributes of a molten globule state, a compact non-native state formed by some proteins, including cyt *c*, under mildly denaturing equilibrium conditions (Kuwajima et al., 1976; Dolgikh et al., 1981; Ptitsyn, 1987; Kuwajima, 1989; Baum et al., 1989; Jeng et al., 1990; Kuroda et al., 1992). Transient folding intermediates with similar properties have also been described on the basis of kinetic CD experiments, gel exclusion chromatography, and transient binding of ANS (Kuwajima, 1989; Ptitsyn et al., 1990; Goldberg et al., 1990). The existence of early folding intermediates with high levels of secondary structure is well documented on the basis of previous kinetic CD studies (Ikeguchi et al., 1986; Kuwajima et al., 1987; Kuwajima, 1991), but there is little direct evidence for chain condensation at early folding stages. Some long-lived intermediates have been shown to be compact by gel exclusion chromatography, and kinetic ANS binding studies provided evidence for the rapid formation of hydrophobic clusters (Ptitsyn et al., 1990).

*The Fast Phase (4–100 ms).* Following the burst phase, the far-UV CD signal exhibits an initial kinetic phase with a time constant of 40 ms. During this fast phase, the ellipticity becomes about 30% more negative, indicating further formation of secondary structure. The transient CD spectrum recorded at the end of the fast phase (100-ms trace in Figure 2) is predominantly helical in character and is consistent with an ~70% contribution from a nativelike CD spectrum. Comparison of the far-UV CD data with the time course of protection for N- and C-terminal amide protons and the quenching of Trp 59 fluorescence (Figure 3) shows that all three probes undergo major changes in the time range from 4 to 100 ms. In contrast, the CD signal in the aromatic region begins to appear only after about 100 ms, and all amide probes outside the N- and C-terminal helices remain largely unprotected for the first 100 ms of refolding.

The present CD and fluorescence results provide additional information on the structural properties of the folding intermediate(s) populated after the fast phase. At 100 ms, the pulse labeling data show partial protection (60%) for amide probes in two helices, which together comprise about 50% of the helical residues of the native cyt *c* structure. On the basis of amide protection, one would thus expect only about 30% of the native helical structure at 100 ms, whereas the peptide CD signal has reached already about 70% of the native level. Apparently, there are contributions to the ellipticity at 222 nm from other helices, such as the 60's and 70's helices, or helices not found in the native structure, which lack sufficiently stable hydrogen bonds to provide measurable protection against exchange. The formation of the early intermediate is also accompanied by a 30% increase in fluorescence quenching, indicative of a further shortening of the average Trp 59–heme distance. However, the absence of an aromatic CD band at 100 ms shows that this reduction in overall chain dimensions does not lead to formation of a tightly packed core with buried aromatic side chains. This conclusion is supported by the observation that amide protons involved in irregular long-range hydrogen bonds (e.g., the hydrogen bond between Trp 59 indole NH and the heme propionate) are still exposed to the solvent at this stage (Roder et al., 1988), suggesting that the intermediate lacks stable tertiary interactions.

*Intermediate and Slow Folding Phases.* All of the experimental probes summarized in Figure 3 exhibit an intermediate folding phase with time constants between 200 and 500 ms and variable relative amplitudes (Table I). The ellipticity at 222 nm reaches the native baseline at about 2 s, indicating that the intermediate phase results in molecules with native-like secondary structure.

The intermediate phase is especially prominent in the kinetics monitored at 289 nm, where 70% of the aromatic CD band appears in a 400-ms phase. Since a negative band in this region of the CD spectrum is caused by aromatic chromophores in the densely packed core of a protein, the ellipticity change indicates the burial of aromatic side chains in the protein interior, or it reflects the stabilization of a loosely packed core, resulting in reduced internal mobility. The presence of a condensed globular structure after the intermediate phase is further supported by the fluorescence of Trp 59, which is almost completely quenched at this stage. On the same time scale, all helical amide probes, including those in the 60's and 70's acquire partial protection, indicating further stabilization of the secondary structure, and some non-helical amide probes witness the formation of tertiary interactions during the intermediate phase (Roder et al., 1988).

The remaining changes for all conformational probes occur in a final phase with a time constant of 5–10 s. This slow phase is particularly pronounced for the aromatic CD signal (27%) and all amide protons outside the N- and C-terminal helices (49%). The time scale of the slow phase is consistent with proline isomerization being the rate-limiting step for the final structural rearrangements that lead to the native state. The involvement of slow-folding species containing non-native proline isomers at this stage was confirmed by pulse labeling experiments under variable labeling conditions (Elöve & Roder, 1991), which revealed the presence of a heterogeneous mixture of fully folded and partially folded molecules in which some of the helices (60's and 70's) are unstable.

Comparison of the relative amplitude for the various experimental probes prior to the slow phase (Figure 3) provides further insight into the structure of the slow-folding species. At about 1 s, the ellipticity at 222 nm is within 5% of the native level and the fluorescence of Trp 59 is ~95% quenched, but the aromatic CD band has only 70% of its native intensity and the protection of some helical amide protons (60's and

70's) as well as nonhelical amide probes is incomplete. Thus, the slow-folding intermediates are compact and have native levels of CD-detected secondary structure, but the packing of the core residues is incomplete, and many hydrogen bonds, including those in the 60's and 70's helices and others involved in tertiary interactions, are not yet fully stabilized.

*Comparison with Previous Kinetic CD Studies.* In their earlier investigation of the cyt *c* folding reaction by kinetic CD, Kuwajima et al. (1987) observed a major burst phase but no significant kinetic changes in the peptide CD signal (less than 20%) on the experimentally accessible time scale, in contrast to the present CD results and the pulse labeling studies. This apparent discrepancy may in part be due to the longer dead time of the earlier measurements (18 ms, compared to 4 ms), after which much of the fast-phase changes have already occurred. Furthermore, the data presented by Kuwajima et al. (1987) were recorded at higher temperature (25 °C, compared to 10 °C) and lower final GuHCl concentration (0.4 M, compared to 0.7 M), conditions which are expected to substantially accelerate the folding process (experiments at 4.5 °C are also mentioned without presenting data). In fact, fluorescence-detected folding experiments at variable final GuHCl concentrations (Elöve, Bhuyan, Luck and Roder, unpublished results) showed a dramatic increase in "missing" amplitude with decreasing denaturant concentration.

The apparent contradiction between the cyt *c* folding kinetics observed by pulse labeling (Roder et al., 1988) and kinetic CD (Kuwajima et al., 1987) was resolved by carefully matching the refolding conditions for the two types of experiments. Thus, the partial helix formation as evidenced by the protection of N- and C-terminal amide protons is indeed accompanied by parallel changes in the far-UV CD spectrum occurring at comparable rates (Figure 3).

*Partially Folded Intermediates Contain Elements of the Native Structure.* Although time-resolved CD spectroscopy has, in principle, the potential for detecting non-native secondary structure during refolding, there is no compelling evidence for this in any of the cyt *c* folding intermediates investigated. In spite of possible contributions from aromatic residues and from the heme to the far-UV CD spectrum, the transient CD spectra can be fitted by a linear combination of the spectra corresponding to native and fully unfolded cyt *c* (Figure 2) at every stage of refolding, suggesting that all intermediates contain only nativelike elements of secondary structure. However, this test is not very sensitive, since $\alpha$-helices make the dominant contribution to the CD spectrum in the region observed (218–235 nm), and a very large fraction of $\beta$-sheet structure would be required to have a detectable effect on the spectrum in this region.

The predominance of nativelike structural elements in folding intermediates is supported by a large body of experimental observations [reviewed by Kim and Baldwin (1990) and Baldwin and Roder (1991)]. For example, only nativelike patterns of protected amide protons were observed at intermediate stages of refolding for ribonuclease (Udgaonkar & Baldwin, 1988, 1990), cyt *c* (Roder et al., 1988; Elöve & Roder, 1991), barnase (Bycroft et al., 1990), and ubiquitin (Briggs & Roder, 1992) and similarly for partially folded acid-denatured forms of apomyoglobin (Hughson et al., 1990) and cyt *c* (Jeng et al., 1990). Other striking demonstrations of the importance of nativelike interactions in folding resulted from recent work on the oxidative folding pathway of BPTI, both for peptide models (Oas & Kim, 1988; Staley & Kim, 1990) and for trapped disulfide intermediates (Weissman & Kim, 1991). On the other hand, evidence for early intermediates with non-native secondary structure, based on kinetic CD spectroscopy, has been reported for some proteins (Kuwajima et al,. 1987, 1991; Chaffotte et al., 1992).

## REFERENCES


Babul, J., & Stellwagen, E. (1972) *Biochemistry 11*, 1195–1200.

Baldwin, R. L., & Roder, H. (1991) *Curr. Biol. 1*, 218–220.

Baum, J., Dobson, C. M., Evans, P. A., & Hanley, C. (1989) *Biochemistry 28*, 7–13.

Blond, S., & Goldberg, M. E. (1986) *Proteins 1*, 247–255.

Blond-Elguindi, S., & Goldberg, M. E. (1990) *Biochemistry 29*, 2409–2417.

Brems, D. N., & Stellwagen, E. (1983) *J. Biol. Chem. 258*, 3655–3660.

Briggs, M. S., & Roder, H. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 2017–2021.

Bushnell, G. W., Louie, G. V., & Brayer, G. D. (1990) *J. Mol. Biol. 214*, 585–595.

Bycroft, M., Matouschek, A., Kellis, J. T., Jr., Serrano, L., & Fersht, A. R. (1990) *Nature 346*, 488–490.

Chaffotte, A. F., Cadieux, C., Guillou, Y., & Goldberg, M. E. (1992) *Biochemistry 31*, 4303–4308.

Creighton, T. E. (1990) *Biochem. J. 270*, 1–16.

Dolgikh, D. A., Gilmanshin, R. J., Brazhnikov, E. V., Bychkova, V. E., Semisotnov, G. V., Venyaminov, S. Y., & Ptitsyn, O. B. (1981) *FEBS Lett. 136*, 311–315.

Elöve, G. A., & Roder, H. (1991) in *Protein Refolding* (Georgiou, G., Ed.) ACS Symposium Series, No. 470, pp 50–63, American Chemical Society, Washington, DC.

Feng, Y., Roder, H., Englander, S. W., Wand, A. J., & DiStefano, D. L. (1989) *Biochemistry 28*, 195–203.

Feng, Y., Roder, H., & Englander, S. W. (1990) *Biophys. J. 57*, 15–22.

Fisher, W. R., Tanuichi, H., & Anfinsen, C. B. (1973) *J. Biol. Chem. 248*, 3188–3195.

Förster, T. (1948) *Ann. Phys. (Leipzig) 2*, 55–75.

Garvey, E. P., Swank, J., & Matthews, C. R. (1989) *Proteins 6*, 259–266.

Goldberg, M. E., Semisotnov, G. V., Friguet, B., Kuwajima, K., Ptitsyn, O. B., & Sugai, S. (1990) *FEBS Lett. 263*, 51–56.

Hughson, F. M., Wright, P. E., & Baldwin, R. L. (1990) *Science 249*, 1544–1548.

Ikai, A., & Tanford, C. (1973) *J. Mol. Biol. 73*, 145–164.

Ikeguchi, M., Kuwajima, K., Mitani, M., & Sugai, S. (1986) *Biochemistry 25*, 6065–6972.

Jeng, M.-F., & Englander, S. W. (1991) *J. Mol. Biol. 221*, 1045–1061.

Jeng, M.-F., Englander, S. W., Elöve, G. A., Wand, A. J., & Roder, H. (1990) *Biochemistry 29*, 10433–10437.

Kim, P. S., & Baldwin, R. L. (1982) *Annu. Rev. Biochem. 51*, 459–489.

Kim, P. S., & Baldwin, R. L. (1990) *Annu. Rev. Biochem. 59*, 631–660.

Knapp, J. A., & Pace, C. N. (1974) *Biochemistry 13*, 1289–1294.

Kuroda, Y., Kidokoro, S., & Wada, A. (1992) *J. Mol. Biol. 223*, 1139–1153.

Kuwajima, K. (1989) *Proteins 6*, 87–103.

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

Kuwajima, K., Yamaya, H., Miwa, S., Sugai, S., & Nagamura, T. (1987) *FEBS Lett. 221*, 115–118.

Kuwajima, K., Garvey, E. P., Finn, B. E., Matthews, C. R., & Sugai, S. (1991) *Biochemistry 30*, 7693–7703.

Labhardt, A. M. (1986) *Methods Enzymol. 131*, 126–135.

Manning, M. C., & Woody, R. W. (1989) *Biochemistry 28*, 8609–8613.

Matthews, C. R. (1991) *Curr. Opin. Struct. Biol. 1*, 28–35.

Miranker, A., Radford, S. E., Karplus, M., & Dobson, C. M. (1991) *Nature 349*, 633–636.
Murry-Brelier, A., & Goldberg, M. E. (1988) *Biochemistry 27*, 7633–7640.
Muthukrishnan, K., & Nall, B. T. (1991) *Biochemistry 30*, 4706–4710.
Oas, T. G., & Kim, P. S. (1988) *Nature 336*, 42–48.
Pace, C. N. (1986) *Methods Enzymol. 131*, 266–280.
Pflumm, M., Luchins, J., & Beychok, S. (1986) *Methods Enzymol. 130*, 519–534.
Ptitsyn, O. B. (1987) *J. Protein Chem. 6*, 273–293.
Ptitsyn, O. B., Pain, R. H., Semisotonov, G. V., Tevovnik, E., & Razgulyaev, O. J. (1990) *FEBS Lett. 262*, 20–24.
Ridge, J. A., Baldwin, R. L., & Labhardt, A. M. (1981) *Biochemistry 20*, 1622–1630.
Roder, H., & Wüthrich, K. (1986) *Proteins 1*, 34–42.
Roder, H., Elöve, G. A., & Englander, S. W. (1988) *Nature 35*, 700–704.
Sears, D. W., & Beychok, S. (1973) in *Physical Principles and Techniques of Protein Chemistry* (Leach, S. J., Ed.) Part C, pp 445–593, Academic Press, New York.
Staley, J. P., & Kim, P. S. (1990) *Nature 344*, 685–688.
Tsong, T. Y. (1974) *J. Biol. Chem. 249*, 1988–1990.
Tsong, T. Y. (1975) *Biochemistry 14*, 1542–1547.
Tsong, T. Y. (1976) *Biochemistry 15*, 5467–5473.
Udgaonkar, J. E., & Baldwin, R. L. (1988) *Nature 35*, 694–699.
Udgaonkar, J. E., & Baldwin, R. L. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 8197–8201.
Wand, A. J., DiStefano, D. L., Feng, Y., Roder, H., & Englander, S. W. (1989) *Biochemistry 28*, 186–194.
Weissmann, J. C., & Kim, P. S. (1991) *Science 253*, 1386–1393.