Article

The Earliest Events in Protein Folding: A Structural Requirement for Ultrafast Folding in Cytochrome *c*

Eefei Chen, Robert A. Goldbeck, and David S. Kliger

J. Am. Chem. Soc., **2004**, 126 (36), 11175-11181• DOI: 10.1021/ja0400077 • Publication Date (Web): 19 August 2004

J|A|C|S
ARTICLES
Published on Web 08/19/2004

The Earliest Events in Protein Folding: A Structural Requirement for Ultrafast Folding in Cytochrome *c*

Eefei Chen,* Robert A. Goldbeck, and David S. Kliger

Contribution from the Department of Chemistry & Biochemistry, University of California, Santa Cruz, California 95064

Received January 7, 2004; E-mail: chen@chemistry.ucsc.edu

Abstract: The folding dynamics of reduced cytochrome *c* (redcyt *c*) obtained from tuna heart, which contains a tryptophan residue at the site occupied by His33 in horse heart cytochrome *c*, was studied using nanosecond time-resolved optical rotatory dispersion spectroscopy. As observed previously for horse heart redcyt *c*, two time regimes were observed for secondary structure formation in tuna redcyt *c*: a fast (microseconds) and a slow (milliseconds) phase. However, the fast phase of tuna redcyt *c* folding was much slower and smaller in amplitude than the same phase in horse. The differences in the fast folding phases suggest that for horse heart redcyt *c*, the conformers that undergo the fastest observed folding have the His18−Fe−His33 heme configuration, which appears to be necessary, but not sufficient, to poise an unfolded chain conformation for fastest folding in redcyt *c*.

Introduction

A recent study using far-UV time-resolved optical rotatory dispersion (TRORD) spectroscopy to follow the folding dynamics after photoinduced reduction of denaturant-unfolded horse heart ferricytochrome *c* observed what may be the fastest folding phase yet detected in this protein, the formation of helical secondary structure within several hundred nanoseconds.[1] The suggestion that this ultrafast ORD signal arises from a small fraction of the total ensemble of unfolded protein chains that is somehow poised for rapid formation of secondary structure motivated the current effort to further understand the origins of this kinetic heterogeneity. That a fraction of the unfolded chains can be distinguished kinetically from the total ensemble suggests not only that its helical folding reaction rate is faster, but that its equilibration with other conformations in the ensemble is slow. Thus, those results add to increasing experimental evidence that kinetic heterogeneity arising from incomplete equilibration of polypeptide chain conformers may be intimately linked with the earliest protein folding events, although it may not affect the time course of slower processes such as those conveniently studied with denaturant mixing techniques.

There is now a large folding kinetics literature, mainly comprising stopped-flow denaturant mixing studies, that shows many proteins fold on time scales of milliseconds or longer through reasonably well-defined transition states that seem to incorporate a significant amount of native state structure.[2] These transition states appear to be in rapid equilibrium with the unfolded states, an observation that implies that the classical transition state theory (TST) developed for small molecules can be applied to folding. Although the possibly heterogeneous dynamics of the unfolded protein chains do not affect folding kinetics in the classical TST view, this possibility has motivated a different, "new view" of folding that describes the conformational heterogeneity by a folding funnel on a rough energy landscape.[3,4] Diffusion around the periphery of the funnel represents interconversion between the unfolded conformations and travel down the funnel represents the free energy-biased diffusion driving those conformations toward the lower configurational entropy of the native state. The conformational diffusion time governing the equilibration of the unfolded chains is pivotal from this point of view. If it is not much faster than the mean time for progress down the funnel to the native state, then the unfolded chain dynamics could affect folding kinetics by inducing, for instance, heterogeneous kinetics (different unfolded ensembles folding in kinetic isolation from one another) or nonexponential kinetics (conformational diffusion becoming rate limiting). On the other hand, if the unfolded chain dynamics are much faster than net progress to the folded state, then the unfolded chains would have time to equilibrate before passage over any more global free-energy barrier, resulting in kinetics that would appear to follow a classical TST picture.

Such early time kinetic heterogeneity would be fundamentally different in nature than the late time (>1 ms) heterogeneity known to arise from certain very specific interactions (e.g., cofactor misligation, non-native disulfide bonds, prolyl peptide bond isomers) that can present high enthalpic barriers near the bottom of a folding funnel. In particular, heterogeneous His33/26 heme ligation and prolyl peptide bond isomerization have been associated with kinetic heterogeneity observed previously in the milliseconds to seconds folding reactions of oxcyt *c*.[5−7]

(1) Chen, E.; Goldbeck, R. A.; Kliger, D. S. *J. Phys. Chem. A* **2003**, *107*, 8149−8155.
(2) Fersht, A. *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding*; W. H. Freeman: New York, 1999.
(3) Bryngelson, J. D.; Wolynes, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 7524−7528.
(4) Bryngelson, J. D.; Wolynes, P. G. *J Phys. Chem.* **1989**, *93*, 6902−6915.
(5) Babul, J.; Stellwagen, E. *Biopolymers* **1971**, *10*, 2359−2361.

10.1021/ja0400077 CCC: $27.50

However, it is the finite speed of the conformational dynamics of the unfolded chains in general, behaving as diffusing heteropolymers, that underlies the ubiquitous potential for kinetic heterogeneity that exists on sufficiently short time scales.

How much time is required for the unfolded chains to equilibrate in protein folding is an open question. However, recent experimental evidence appears to bracket this time between the submicrosecond time scales accessible to the computer simulations that have provided impetus to the funnel-landscape view and the millisecond and longer time scales typically investigated in experimental kinetic studies. Kinetic heterogeneity observed for the earliest events in horse heart redcyt *c* folding suggests that the conformational diffusion time constant, $\tau_{\text{diffusion}}$, is on the order of ~10−100 μs (diffusion may proceed over several time constants before heterogeneity is reduced to undetectable levels, depending on the sensitivity of the probing technique).[1,8] Similarly, a study of a λ repressor fragment (λ_{6-85}) wild type and several mutants showed that for the fastest folding mutants the rate coefficient behaved nonexponentially below 4 μs, suggesting disequilibrium between the unfolded and transition states.[9] An equilibration time constant of ~2 μs was inferred for λ_{6-85} from this nonexponential behavior. That observation may be reasonably consistent with the range inferred previously for redcyt *c*, particularly given that both the shorter chain length and the significantly higher relaxation temperature for the λ repressor fragment study would be expected to shorten the equilibration time. However, cross-linking of the chains by prosthetic group coordination in redcyt *c* complicates comparison. The evidence from both proteins suggests that slow conformational diffusion can affect folding kinetics, at least on microsecond and shorter time scales, by impeding equilibration among the unfolded chains.

In addition to the requirement that conformational equilibration be slow enough in denatured proteins to introduce kinetic heterogeneity into their fast folding reactions, the other basic requirement for the observation of such kinetic heterogeneity is that kinetically isolated unfolded ensembles have different reaction rate constants (or at least mean progress rates if TST is not valid) for folding or some other detectable process. We present here an attempt to correlate the kinetic heterogeneity of these putative conformational ensembles to specific protein structural features by a species comparison of submillisecond ("burst phase" of denaturant mixing experiments) secondary structure formation in redcyt *c* from horse and tuna heart. Besides introducing conformational heterogeneity to the bulk protein structure by removing most organized secondary and tertiary structure, the addition of denaturant such as the guanidine hydrochloride (GuHCl) used in this study also promotes in redcyt *c* the replacement of the native ligand Met80 with non-native ligands at the heme prosthetic group's axial site. These non-native ligands have been identified largely with histidine residues, and in GuHCl concentrations above 1.5 M the bis-His [His18-Fe(III)-His33/26] form is the dominant ferric heme ligation.[5−7,10] To examine the possible role of His33/26 in the early-time kinetic heterogeneity observed in horse heart redcyt *c* folding,[1,8] we studied the dynamics of redcyt *c* folding in tuna heart wherein His33 is replaced by a tryptophan residue. Although the structures of the horse heart and tuna heart cytochrome *c* species are very similar, in keeping with their ~80% sequence identity, the dynamics of the fast folding phase (<1 ms) for the two species are dramatically different. These results suggest that the extraordinarily rapid formation of secondary structure in horse redcyt *c* is crucially dependent on the presence of a single residue, His33.

Materials and Methods

Sample Preparation. Tuna heart-oxidized cytochrome *c* (oxcyt *c*) and NADH (Sigma), sodium phosphates (NaP: monobasic, Na_2HPO_4, and dibasic, NaH_2PO_4) (Fisher Scientific), sodium hydrosulfite (dithionite) (Fluka), and ultrapure GuHCl (ICN Biochemicals) were purchased and used without further purification. Samples for TRORD experiments were prepared in buffer solutions containing 0.1 M NaP and either 3.3 or 4 M GuHCl to obtain a final protein concentration of ~70 μM (pH 7). Before deoxygenating the protein solution with nitrogen gas for 15−30 min, it was stirred in air for about 15 min to oxidize any traces of redcyt *c*. The oxygen-free sample, as well as a separate, deoxygenated vial containing solid NADH (~500 μM), was placed in a glovebag continuously purged with nitrogen gas. The oxcyt *c* sample was then added to the solid NADH in the glovebag.

Tuna oxcyt *c* samples for the equilibrium ORD measurements used to determine the folding curves were prepared by titrating a solution of oxcyt *c* (~60 μM) in 0.1 M NaP and 6.1 M GuHCl (pH 7) with a second solution of oxcyt *c* (~60 μM) in 0.1 M NaP (pH 7) to various concentrations of GuHCl. The pH of all samples was adjusted to 7, after which the oxcyt *c* samples were deoxygenated and placed in a glovebag purged with nitrogen. Redcyt *c* samples were prepared by adding sodium dithionite to the oxcyt *c* samples. For all sample preparations, the GuHCl concentrations in the pure buffer and in the oxcyt *c* solutions were measured on an ABBE-3L refractometer (Milton Roy).

Time-Resolved ORD Experiments. A ~15 mJ, 355-nm pulse (7 ns, full-width half-maximum) from a Quanta Ray DCR-1 Nd:YAG laser was used to initiate rapid oxidation of NADH and the subsequent reduction of oxcyt *c*. The resulting time-dependent spectral changes were probed with a time-resolved ORD apparatus. The pump beam (1 × 2 mm) entered the sample cell at an angle of 15° to overlap with the probe beam (300 μm), whose propagation axis was normal to the face of the sample, from a xenon flash lamp. Initially unpolarized, the probe beam passed through a MgF_2 polarizer, the sample, and a second polarizer. Subsequently, the probe beam was focused onto the slit of a spectrograph, dispersed with a 600 grooves/mm grating (200 nm blaze), and then detected by a gated optical multichannel analyzer (Princeton Applied Research). The probe beam was focused down from a 6-mm diameter by using two mirrors positioned around the sample, after and before the first and second polarizers, respectively. ORD measurements were made by rotating the first polarizer by angles of $+\beta$ and $-\beta$ from the crossed (90°) position of the two polarizers, where $\beta = 1.87°$ in these studies. The ORD signal is equal to the difference between the signals measured at $+\beta$ and $-\beta$ divided by their sum. Further details about the TRORD method can be found in previously published literature.[11]

TRORD data for redcyt *c* in 3.3 and 4 M GuHCl were obtained at 27 time delays from 270 ns to 800 ms and from 270 ns to 1 s, respectively, after initiation of folding. A total of 640 averages were

(6) Elöve, G. A.; Bhuyan, A. K.; Roder, H. *Biochemistry* **1994**, *33*, 6925−6935.
(7) Colón, W.; Wakem, L. P.; Sherman, F.; Roder, H. *Biochemistry* **1997**, *36*, 12535−12541.
(8) Goldbeck, R. A.; Thomas, Y. G.; Chen, E.; Esquerra, R. M.; Kliger, D. S. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 2782−2787.
(9) Yang, W. Y.; Gruebele, M. *Nature* **2003**, *423*, 193−197.
(10) Thomas, Y. G.; Goldbeck, R. A.; Kliger, D. S. *Biopolymers (Biospectroscopy)* **2000** *57*, 29−36.
(11) Shapiro, D. B.; Goldbeck, R. A.; Che, D.; Esquerra, R. M.; Paquette, S. J.; Kliger, D. S. *Biophys. J.* **1995**, *68*, 326−334.

accumulated at each time delay from 270 ns to 2 μs, and 384−608 averages were acquired at all time delays between 3 μs and 800 ms for redcyt *c* folding in 3.3 M GuHCl. For folding in 4 M GuHCl, 512 averages were collected for time delays from 270 ns to 2 μs and 260 averages were accumulated at each time delay from 3 μs to 1 s. The equilibrium oxcyt *c* and redcyt *c* ORD signals were each determined from about 704 averages.

A peristaltic pump in the glovebag recycled the sample through a 1.5-mm path length flow cell with fused silica windows and then back to a collection vial inside the bag. The UV−vis spectrum (Shimadzu UV-2101PC spectrophotometer) of the collected solution was compared to the spectrum of fresh starting material before the sample was cycled through additional experiments. Data were collected at a repetition rate of 0.5 Hz, with the sample flowed at a rate of 8 μL/s, between laser pulses only. This flow rate moved the volume of irradiated sample out of the pump−probe path before the next laser pulse, minimizing accumulation of photoproduct or photodegradation products. The GuHCl concentrations and the pH of the oxcyt *c* solutions were measured before and after the experiments. Sample temperature (24 to 25 °C) was checked frequently during the experiment with a Fluke 80T-IR infrared temperature probe.

Equilibrium ORD Experiments. The equilibrium ORD data used to calculate the folding curves for oxcyt *c* and redcyt *c* were also measured on the TRORD apparatus. After preparation, as discussed above, all solutions were handled in a nitrogen-purged glovebag. Using a peristaltic pump, we flowed the oxcyt *c* solution to the 1.5-mm path length flow cell for data acquisition and then returned the solution to the glovebag, where dithionite was added to generate redcyt *c*. A portion of the redcyt *c* sample was cycled back to the cell for ORD measurements, and the remaining solution was used to obtain a UV−vis spectrum. A UV−vis spectrum of the corresponding oxcyt *c* was also recorded. The folding curves for horse cyt *c* were obtained from previously reported CD data.[10] At all concentrations of GuHCl, 192−256 averages were accumulated for each of the oxcyt *c* and redcyt *c* ORD signals.

Data Analysis. The TRORD data were analyzed in the form of difference kinetic traces, obtained by averaging the difference ORD signals (time-resolved spectra minus initial oxcyt *c* spectrum) from 230 to 234 nm. Exponential fitting analyses, written in the mathematical software package Matlab (The MathWorks), were performed on these kinetic traces. To fairly compare the current study with the TRORD results on horse redcyt *c*, the data were also analyzed with the three longest time measurements removed from the kinetic fitting.[1] This is because at long delay times (hundreds of milliseconds) reoxidation due to the presence of trace oxygen competes with protein folding. Percentages of secondary structure reported were calculated relative to the ORD signal for the fully reduced, folded protein.

Data points on the tuna folding titration curves were generated by calculating the fraction of unfolded protein from the intensities of the ORD signals at 232 nm (data averaged from 230 to 234 nm) for each GuHCl concentration. The values at 232 nm for 0 and 6 M GuHCl were used to represent the fully folded and "fully unfolded" protein.

Results

Figure 1 shows the changes in the fraction of unfolded secondary structure present in solution as a function of GuHCl concentration for tuna oxcyt *c* and redcyt *c*. The data points on the tuna folding curves, obtained from ORD measurements, are overlaid with the folding curves for horse oxcyt *c* and redcyt *c*, which were calculated from CD data and previously presented in Thomas et al.[10] The folding curves indicate that in the presence of denaturant the respective oxcyt *c* and redcyt *c* for horse and tuna have similar folding free energies. These results

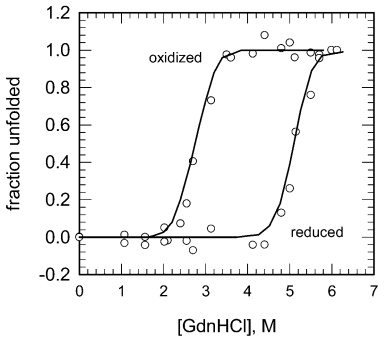

Figure 1. Comparison of GuHCl denaturant titrations of horse heart and tuna heart cyt *c*. The folding curves for horse (black lines) and tuna (circles) oxcyt *c* and redcyt *c* can be overlaid, demonstrating that the folding free energies for the respective oxidation states are very similar. ORD measurements (1.5-mm path length cell) were used to calculate the folding curves for tuna cyt *c*. The curves for horse cyt *c* are taken from the CD data of Thomas et al.[10] The tuna cyt *c* samples (~60 μM) were prepared in 0.1 M NaP at pH 7.

are consistent with previous equilibrium and kinetic studies that indicate similar unfolding free energies for horse and tuna cyt *c*.[12]

It was demonstrated previously that TRORD and TRCD spectroscopy yield similar kinetic results for the horse heart redcyt *c* system.[1] ORD is preferred in folding studies of redcyt *c* because it offers a signal-to-noise advantage over CD experiments, particularly under conditions where high sample absorbance leads to low light intensities. The CD spectrum (208 and 222 nm) that is characteristic of the n $\rightarrow \pi^*$ peptide transitions has a corresponding negative ORD signal at 232 nm. The equilibrium ORD signals for oxcyt *c* and redcyt *c* are shown in Figure 2.

Figure 2 also shows the kinetic behavior of the ORD signals for tuna redcyt *c* in 3.3 M (A) and 4 M (B) GuHCl following rapid photoreduction of oxcyt *c*. Although redcyt *c* folding was probed at 27 time delays, only three signals, measured at 1, 100, and 500 ms, are shown to demonstrate the trend in the signal over time. All measured time-dependent data are represented in the kinetic traces of Figure 3. The time-dependent changes of the ORD signal are assigned to the formation of secondary structure because of the similarity in band shape and sign to that for redcyt *c*. As with studies of redcyt *c* from horse, we discuss tuna folding dynamics in terms of two major time regimes: a fast (270 ns to 1 ms) and a slow (>1 ms) regime. The fast phase correlates with the unobserved "burst" phase of earlier studies in which time resolution was typically limited by the dead time of denaturant dilution techniques, whereas the slow phase is characterized by kinetics on the hundreds of millisecond time scale.

The kinetic traces for tuna redcyt *c* folding (Figure 3) do not show the ultrafast formation of secondary structure that in part characterizes the unusual behavior of the fast phase folding dynamics for horse redcyt *c*. The differences in the fast folding kinetic phases of horse versus tuna redcyt *c* are summarized in Figure 4. A large increase in secondary structure was detected within several hundred nanoseconds for the folding dynamics of horse redcyt *c* in 4 M GuHCl.[1] In addition, the rise time of secondary structure formation in the horse data slows significantly from 4 M ($\tau \leq 0.4$ μs) to 3.3 M ($\tau \approx 2$ μs). This is

(12) McLendon, G.; Smith, M. *J. Biol. Chem.* **1978**, *253*, 4004−4008.

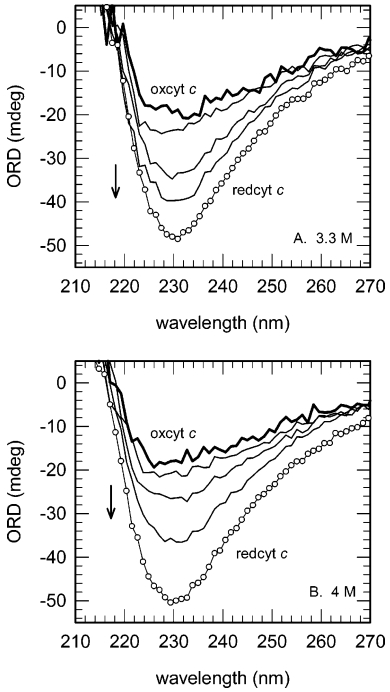

Figure 2. TRORD signals for tuna redcyt *c* folding triggered by photoinduced electron transfer from NADH. The ORD signals measured at 1, 100, and 500 ms (thin black line) after initiation of folding in (A) 3.3 M and (B) 4 M GuHCl are shown along with the signals for equilibrium oxcyt *c* (thick black line) and redcyt *c* (circles). Only three of the 27 measured time-delayed signals are shown to more clearly represent the major changes in the ORD signals over time.

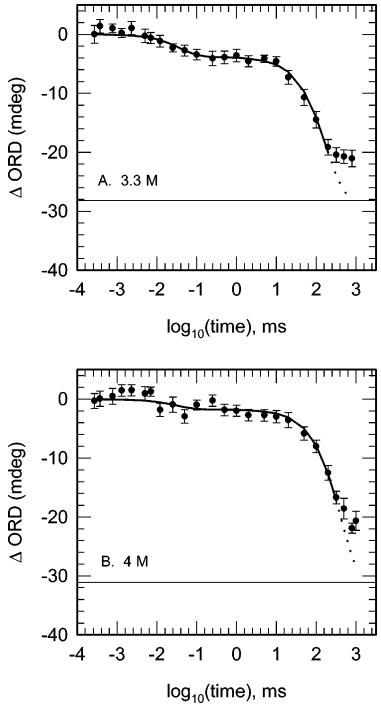

Figure 3. Kinetics of tuna redcyt *c* folding triggered by photoinduced electron transfer from NADH. The difference ORD signals for folding in (A) 3.3 M and (B) 4 M GuHCl are shown along with respective two-exponential fits to the data. The fit in the region of the last three time delays, which were not included in the analysis, is represented with a dotted line. The respective equilibrium redcyt *c* ΔORD signals are denoted with thin black lines. The initial denatured sample, ~70 μM oxcyt *c*, was prepared in 0.1 M NaP (pH 7) with 500 μM NADH. All data were collected at 25 °C in a 1.5-mm path length cell.

clearly not the case for fast phase folding of tuna redcyt *c* in 3.3 and 4 M GuHCl. The TRORD signal begins to change at ~5 μs, showing a steady increase in magnitude until ~100 μs, at which point the intensity of the signal is about 14 and 6% of that for equilibrium redcyt *c* in 3.3 and 4 M GuHCl, respectively. There is then no significant change in the ORD signal until after 10 ms. The exponential time constants (fractional amplitude) associated with secondary structure formation during this fast phase are 40 μs (0.14) and 28 μs (0.08) for folding in 3.3 and 4 M GuHCl, respectively. The fractional amplitude of this component decreases by a factor of ~1.8 with higher concentrations of GuHCl and closely reflects the percentage of secondary structure that has formed during the fast phase (<1 ms). This contrasts with the observation from horse redcyt *c* folding in that, although the fractional amplitude decreases by about 1.5 in 2.7 M versus 4 M GuHCl, the amount of secondary structure formed in the fast phase is approximately constant.

The dynamics of the slow phase in tuna redcyt *c* is similar to that observed for horse redcyt *c* and is characterized by a component with a lifetime that increases from 185 ms (0.86) in 3.3 M GuHCl to 430 ms (0.92) in 4 M GuHCl (see Figures 3 and 4). This process has been correlated with native secondary structure formation[13] that results from the displacement of the non-native His ligation by the native Met80 ligand.[6,7] Whereas the time constants show behavior that is consistent with expectations that folding is slower in higher concentrations of denaturant,[14,15] the amplitude for this process also increases with the concentration of GuHCl. Because fully folded redcyt *c* is expected from the reaction thermodynamics to form essentially quantitatively at the denaturant concentrations studied, the amplitude increase was not due to formation of more secondary structure but, rather, to the presence of less secondary structure in the starting oxcyt *c* sample at higher denaturant concentration.[1,10] For folding observed in both 3.3 and 4 M GuHCl, there was formation at the longest time delay measured (~1 s) of 85 and 70% secondary structure relative to equilibrium redcyt *c*, respectively. Observation of less than the nearly 100% secondary structure formation expected from the folding free energy can be explained here by the competition of reoxidation due to trace levels of adventitious oxygen with the folding reaction of redcyt *c* and by the expectation that higher concentrations of denaturant will increase the time required to reach complete folding. The presence of trace levels of oxygen apparently did not affect the earliest events in protein folding, which are the focus of this study.

Discussion

Fast folding reactions have been triggered previously in redcyt *c* in studies using ligand photodissociation[14] and electron transfer[15−17] methods. These techniques are based on the difference in the folding free energies between CO-ligated redcyt

(13) Chen, E.; Wittung-Stafshede, P.; Kliger, D. S. *J. Am. Chem. Soc.* **1999**, *121*, 3811−3817.
(14) Jones, C. M.; Henry, E. R.; Hu, Y.; Chan, C.-K.; Luck, S. D.; Bhuyan, A.; Roder, H.; Hofrichter, J.; Eaton, W. A. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 11860−11864.
(15) Mines, G. A.; Pascher, T.; Lee, S. C.; Winkler, J. R.; Gray, H. B. *Chem. Biol.* **1996**, *3*, 491−497.
(16) Pascher, T.; Chesick, J. P.; Winkler, J. R.; Gray, H. B. *Science* **1996**, *271*, 1558−1560.
(17) Telford, J. R.; Wittung-Stafshede, P.; Gray, H. B.; Winkler, J. R. *Acc. Chem. Res.* **1998**, *31*, 755−763.

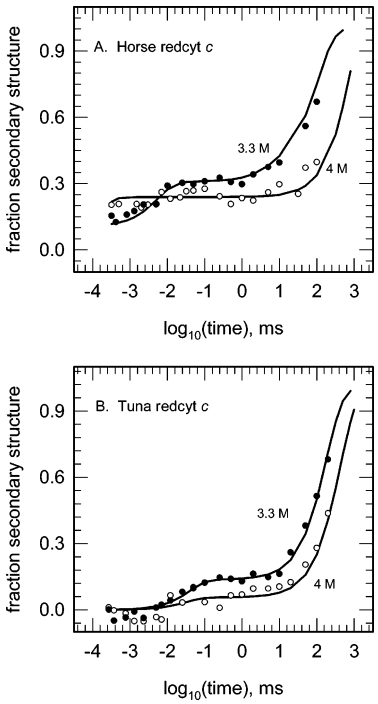

Figure 4. Comparison of (A) horse and (B) tuna redcyt *c* folding kinetics in 3.3 M (black circles) and 4 M GuHCl (white circles). Fractional secondary structure content was calculated by comparing the ΔORD (230 nm) measured at each delay time to the ΔORD of the final folded state. The experimental data points are shown with the two-exponential fits to the data.

c and CO-deligated redcyt *c*, or between oxcyt *c* and redcyt *c*, respectively. In the range of GuHCl concentrations from 3 to 5 M, more than 95% of the redcyt *c* protein will be in the folded configuration, whereas less than 5% of oxcyt *c* will be folded (see Figure 1); hence, rapid reduction of oxcyt *c* triggers folding. In the case of CO photodissociative triggering of horse cyt *c*, ~10% secondary structure is formed within ~1 μs. After this time, non-native His26/33 binding (50−100 μs) and CO recombination (hundreds of microseconds) compete with and frustrate further formation of organized secondary or tertiary structure.[18] In contrast, the rapid photoreduction method used in the present study allowed folding to be probed essentially to completion.

Our finding that tuna and horse cytochromes *c* have very fast (<1 ms) kinetic phases that differ sharply in rate, amplitude, and denaturant dependence despite having very similar native structures, stabilities, and slow kinetic phases for folding strongly suggests that the presence of His33 in the horse sequence is associated with the more rapid and robust submillisecond helix formation observed in that protein. Although roughly 20% of the residues differ between these two species, variations in the histidine residues are expected to have the most profound influence on the dynamics of the unfolded chains via the constraint imposed on chain conformation by non-native ligation to heme iron. His33−Fe−His18 and His26−Fe−His18 constitute 80 and 20%, respectively, of the heme ligation states of unfolded chains in horse ferricytochrome *c* under these conditions.[7] The denatured H33N horse variant, on the other hand, is dominated by His26−Fe−His18 coordination (~80%), the unidentified minor component presumably being due to one or more of the 19 lysine residues or the partially unacetylated N-terminus of the recombinant protein studied by Colón et al. (acetylation prevents the N-terminus of protein expressed in horse from binding to the heme iron). The H33W difference between horse and tuna sequences thus leads us to expect that the unfolded starting material for the present tuna cyt *c* experiments will also be dominated by His26−Fe−His18 ligation. Although three of the lysines present in horse are missing in tuna, that difference is not as likely as the missing histidine to have a significant effect on the unfolded chains because lysine−heme ligation is expected to be at best a minor component under these conditions of temperature and pH.

It is not clear from the present results whether His33−Fe−His18 coordination actually speeds helix formation or just impedes it less than His26−Fe−His18 coordination. Studying the H33/26 double mutant could shed light on this question, where the observation of a slower rate of ultrafast helix formation in the absence of both non-native histidine ligands would support the former scenario. The evidence from CO photodissociation studies, in which an ultrafast helix-folding reaction proceeded from a pentacoordinate heme species (His18-Fe···), is somewhat ambiguous on this point.[8,18] (Note that Goldbeck et al. tentatively assigned that process to native state formation because of a rough correspondence in amplitude and rate observed in far-UV TRCD with those for a Met−Fe−His heme ligation process observed in Soret absorption and magnetic circular dichroism spectroscopies.[8] However, there is no evidence that the ultrafast process observed here in photoreduced horse cyt *c* is different in nature than that for CO photodissociation, and the denaturant dependence of the former, discussed further below, suggests that these processes do not form native state.) On one hand, the fastest folding kinetics observed after CO photodissociation conducted at 4.6 M GuHCl was similar to that for the 4.0 M His33−Fe−His18 photoreduction results in that both time constants are apparently shorter than the instrumental time resolution, ~500 ns.[18] On the other hand, the corresponding amplitude was less than half of that observed in the photoreduction case. The apparent decrease in amplitude in the absence of His33−Fe−His18 coordination suggests that it may promote helix formation compared with no tethering of the protein chain at the sixth heme coordination site, but this comparison assumes that the amplitude of the 4.0 M photoreduction data can be extrapolated to 4.6 M GuHCl. The time constant comparison is also ambiguous in that we cannot say if His33 coordination speeds the process compared with pentacoordinate heme, but it does suggest that His26 coordination impedes helix formation compared with either pentacoordinate or His33-coordinated heme. Altogether, this evidence suggests that the most likely order of the fast phase helix-forming propensities of the three heme coordination types is His33−Fe−His18 > His18−Fe··· > His26−Fe−His18.

The ultrafast TRORD/CD kinetics observed in horse cyt *c* and the slower microsecond kinetics observed here for tuna represent helix formation processes that appear to be faster than all other folding processes reported previously for cyt *c*, including a chain collapse phase generally reported to lie in the 50−100 μs time range.[19,20] These helical processes thus

(18) Chen, E.; Wood, M. J.; Fink, A. L.; Kliger, D. S. *Biochemistry* **1998**, *37*, 5589−5598.
(19) Hagen, S. J.; Eaton, W. A. *J. Mol. Biol.* **2000**, *297*, 781−789.
(20) Shastry, M. C.; Roder, H. *Nat. Struct. Biol.* **1998**, *5*, 385−392.

correspond to the "burst" phase inferred to develop during the ~1-ms dead time of previous denaturant-mixing CD studies of cyt *c*. In classical kinetic terms, these early helix formation processes would be described as U → I, where U is the unfolded state and I is an intermediate, which in this case appears to have native-like amounts of helical secondary structure but little native tertiary structure. However, it is not clear whether this intermediate is obligatory to formation of the native state N, as in a framework mechanism (U ⇄ I → N) wherein early formation of secondary structure provides a framework facilitating formation of native tertiary contacts,[21] or nonproductive (I ⇄ U → N). Regardless of its fate, the formation of this intermediate appears to contain important clues about the dynamics of the unfolded chains as they progress toward conformational equilibrium, complete conformational equilibrium being an implicit assumption underlying classical transition state pictures of folding kinetics.

The ultrafast helical folding phase of horse cyt *c* was remarkable in this regard because it strongly suggested the presence of kinetic heterogeneity in the unfolded chains.[1] A possible origin of this heterogeneity was slow conformational diffusion, inferred previously from TRMCD evidence to affect the heme religation reactions triggered by photodissociation of cyt *c*-CO.[8] Note, however, that this effect does not seem to be as evident in less heme coordination-sensitive absorption data for CO photodissociation.[22] An alternative explanation for the horse photoreduction result was that the fraction of the sample undergoing ultrafast helix formation (~20%) simply corresponded to the minority heme ligation species His26−Fe−His18. In other words, such early-time kinetic heterogeneity might have arisen from heterogeneous heme ligation, interchange of the heme coordination forms being rate-limited by what may be very slow histidine off rates. The further significance of the present results is thus that they appear to rule out this "trivial" explanation for the early-time kinetic heterogeneity observed in horse because that scenario would predict from the dominant His26 coordination of tuna cyt *c* a dramatically larger and faster early folding phase with a very different denaturant dependence for the latter protein than was observed.

The different effects of denaturant on the early helix formation phases of tuna and horse cyt *c* are interesting in that they imply different kinetic mechanisms for these processes. Increasing denaturant concentration often slows the kinetics of reactions that produce natively folded protein by raising the free energy of activation in a linear manner that is proportional to the decrease in driving force. The slow folding phases observed here for tuna and horse cyt *c* both followed this trend as expected since both produce fully folded protein. However, the direction of the effect observed in the ultrafast horse phase was surprising in that helix formation accelerated at high concentration, suggesting that this process does not produce natively folded protein. Because the rate did not follow the expected correlation with driving force, that suggests that denaturant directly perturbs the transition state (or at least the kinetic barrier over which ensembles are diffusing) so as to stabilize it relative to U. This would be surprising in a transition leading to the native state, such a transition state being expected to contain already a great deal of native structure and thus be destabilized by denaturant relative to U. However, the transition state (barrier to diffusion) in question here probably contains little native structure and could plausibly involve some additional unfolding of residual structure present in the (20% populated) subensemble of U out of which this process proceeds. We have little evidence for the specific nature of this residual structure except for the dramatic kinetic effect of replacing His33 with a tryptophan residue inferred from comparing the tuna and horse data. This evidence tends to point to transient structure in U associated with coordination of the protein chain to the heme via His33. Residual structure under denaturing conditions is known to be a feature of many proteins.[23] The residual structure detected by small-angle X-ray scattering in the unfolded state of oxcyt *c*, persisting at GuHCl concentrations up to about 5 M, probably arises from nonspecific hydrophobic interactions and is devoid of stable secondary structure.[24] For the fast phase of tuna cyt *c*, on the other hand, the effect of increased denaturant (a drop in amplitude accompanied by a similar decrease in observed time constant) is more consistent with a fast equilibrium involving a strong back reaction and a transition state with a more typical tendency to destabilization by denaturant. In this case (which would correspond to the ascending arm in a typical chevron plot of log k_{obs} vs denaturant), the observed rate constant, $k_{obs} = k_f + k_b$, tends to be dominated by the back reaction rate constant, k_b, which increases with denaturant.

The absence in tuna redcyt *c* of the very fast (submicrosecond) folding kinetics observed in horse indicates that His18−Fe−His33 heme coordination is a necessary requirement for an unfolded chain conformation to be a member of the putative fast folding ensemble. However, it cannot be sufficient because ~80% of the unfolded chains of horse oxcyt *c* (the starting material in photoreduction-induced folding experiments) possess this heme configuration,[7] whereas a much smaller fraction of the unfolded chains participate in the ultrafast process in horse. The TRORD amplitude and relatively small fraction of sample able to react with photoinduced reductant before a microsecond together implied that this fraction of the sample, ~20%, acquires a near-native amount of secondary structure in that process, rather than the bulk sample acquiring a small amount of secondary structure.[1] Two obvious scenarios could prevent the unfolded chains in a homogeneously equilibrated ensemble of His18−Fe−His33-coordinated conformations from completely converting to the misfolded transient: (1) a strong back reaction might produce a weakly converted equilibrium mixture (as suggested above for the His18−Fe−His26 tuna data) and (2) kinetic competition from a second, CD/ORD-silent process with a faster reaction rate constant might sharply reduce the yield of the observed folding process. Although we cannot absolutely exclude either possibility, they both appear unlikely in light of the relative denaturant independence of the reaction amplitude in the horse protein because (1) it seems very unlikely that increasing denaturant concentration would not shift an equilibrium between unfolded and partially folded states and (2) there

(21) (a) Ptitsyn, O. B. *Dokl. Akad. Nauk SSSR* **1973**, *210*, 1213−1215. (b) Bashford, E.; Cohen, F. E.; Karplus, M.; Kuntz, I. D.; Weaver, D. L. *Proteins: Struct., Funct., Genet.* **1988**, *4*, 211−227. (c) Kim, P. S.; Baldwin, R. L. *Annu. Rev. Biochem.* **1990**, *59*, 631−660. (d) Karplus, M.; Weaver, D. L. *Protein Sci.* **1994**, *3*, 650−668.
(22) Hagen, S. J.; Latypov, R. F.; Dolgikh, D. A.; Roder, H. *Biochemistry* **2002**, *41*, 1372−1380.
(23) Millet, I. S.; Doniach, S.; Plaxco, K. W. *Adv. Protein Chem.* **2002**, *62*, 241−262.
(24) Segel, D. J.; Fink, A. L.; Hodgson, K. O.; Doniach, S. *Biochemistry* **1998**, *37*, 12443−12451.

is no obvious candidate for a competitive CD/ORD-silent process, and moreover, any kinetic competition it presented would have to involve coincidentally identical denaturant effects on the two rate constants. Therefore, this process probably forms native-like amounts of secondary structure in a (20% populated) subensemble of protein chains that react heterogeneously because of slow conformational diffusion, these His33-ligated unfolded chains also having backbone and side chain conformations that poise them for fast helix formation. If equilibration of the unfolded chain ensembles was rapid, then all of the unfolded chains would quickly sample such a conformation and many more would fold as well. The observation that this is not the case for the ultrafast folding process in horse redcyt *c* implies that its unfolded chain conformations are not in facile equilibrium, i.e., that unbiased conformational diffusion around the funnel rim is relatively slow.

It is tempting to speculate that the qualitatively different fast phases of tuna vs horse cyt *c* reflect homogeneous vs heterogeneous kinetic mechanisms, respectively, as the unfolded conformational diffusion time constant may lie between their two helical folding time constants at a given denaturant concentration and allow more complete conformational equilibration in tuna. The time constants observed here for tuna cyt *c* were clearly slower than the average time constant for reduction, 5−10 μs (the latter time constant being somewhat ambiguous because the photoreduction process was non-first order and involved heterogeneous reductants[1]). This tends to remove in tuna the most compelling argument for the presence of early time kinetic heterogeneity in horse cyt *c*: the need to explain the apparently paradoxical appearance of secondary structure at a rate faster than the overall rate of photoreduction triggering that process, in the presence of high denaturant.[1] The second argument for kinetic heterogeneity in horse, the need to account for incomplete conversion of U to I by an apparently irreversible process (U $\rightarrow$ I), is also absent in tuna, wherein the reaction appears to be reversible (U $\rightleftarrows$ I). However, while there is no evidence for fast phase kinetic heterogeneity, and at least some equilibration might be expected from previous estimates of $\tau_{\text{diffusion}}$, neither is there positive evidence that the unfolded conformations of tuna cyt *c* are partially or completely equilibrated on this time scale. Hence, the present results do not clearly distinguish between heterogeneous vs homogeneous kinetics for the early helix formation phase in tuna cyt *c*, and they further establish the time frame for conformational diffusion only indirectly, by ruling out a possible alternative explanation for kinetic heterogeneity in horse cyt *c*.

Finally, why His18−Fe−His33 heme coordination would facilitate very fast helix formation in redcyt *c* is not clear, but the most significant alternative heme configuration, His18−Fe−His26, may be too sterically constrained by the relatively short chain length between the two histidine residues to adopt chain conformations needed for the most facile helix formation. The considerably smaller amplitudes observed for the fast folding phase of tuna redcyt *c* at both denaturant concentrations studied may reflect the smaller driving force for helix formation due to such hindrance (assuming that this phase arises from the His26-ligated ensembles dominant in tuna and not a minor component). Molecular modeling of folding in His33- and His26-ligated proteins may provide useful insights into this question.

Conclusions

This study has focused on understanding the possible structural basis for kinetic heterogeneity in the earliest folding events by comparing the fastest helix formation reactions of tuna and horse heart redcyt *c*. Besides strengthening the case for conformational-diffusion-linked kinetic heterogeneity in photoreduced horse cyt *c* by ruling out heterogeneous heme ligation as an alternative explanation, these results further suggest that His18−Fe−His33 coordination is necessary, but not sufficient, to facilitate the fastest helical folding in redcyt *c*. This coordination appears to be associated with faster helical folding than His18−Fe−His26 coordination and may promote faster folding than the less heme-tethered His18−Fe··· coordination. Moreover, only a small subset of the conformations accessible to unfolded chains possessing His18−Fe−His33 coordination may have the proper orientation of backbone and side chain degrees of freedom additionally needed to facilitate ultrafast folding, a requirement that appears to overlap with the residual ordered structure known to persist in a subset of the unfolded conformers.

Acknowledgment. This study was supported by National Institute of General Medical Sciences (NIH) Grant GM38549.

JA0400077