J|A|C|S COMMUNICATIONS

Published on Web 10/13/2006

Direct and High Resolution Characterization of Cytochrome *c* Equilibrium Folding

Laura B. Sagle, Jörg Zimmermann, Philip E. Dawson, and Floyd E. Romesberg*

Department of Chemistry, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037

Received July 19, 2006; E-mail: floyd@scripps.edu

Understanding how proteins fold has emerged as a central problem in biophysics. The equilibrium folding mechanism of cytochrome *c* (cyt *c*) has served as a model system.[1] Analysis of spectroscopic data,[1c] along with the results of small-angle X-ray scattering,[2] fluorescence,[3] magnetic circular dichroism,[4] amide exchange,[5] 2D IR,[6] and NMR[1c,7] studies convincingly demonstrate that intermediates accumulate. Unfortunately, the characterization of both the folding process and of the intermediates has been limited by the low structural and/or temporal resolution of the available techniques.

We have been developing the use of site-selectively incorporated carbon−deuterium (C−D) bonds as probes of protein folding and dynamics.[8] C−D bonds may be incorporated anywhere within cyt *c*, and their environmentally sensitive IR absorptions are easily characterized. In addition to high structural resolution, this technique has high temporal resolution, owing to the inherently fast time scale of IR spectroscopy. Thus, even the most rapidly interconverting folding intermediates may be observed. Previously, cyt *c* was semisynthesized with either (methyl-d_3) methionine incorporated at Met80 ((d_3)Met80) or with a 1:1 mixture of C_{δ_1}-d_3 and C_{δ_2}-d_3 labeled leucine isotopomers incorporated at position 98 ((d_3)-Leu98).[8c] The IR spectra of (d_3)Met80 revealed the population of a folding intermediate followed by a transition to a solvent exposed environment, while the Leu98-labeled protein showed a simple two state transition. To more fully understand this folding process, we have characterized the spectra of (d_3)Leu68, (d_8)Lys72, (d_8)Lys73, (d_8)Lys79, (d_3)Met80, (d_3)Ala83, (d_3)Leu94, (d_3)Leu98, and (d_3)-Ala101 cyt *c* as a function of GdnHCl and urea (Figure 1). These residues are located within four structural motifs of the folded protein; the 60's helix (residues 60−68), the short 70's helix (residues 70−75),[9] the D-loop (residues 76−84), and the C-terminal helix (residues 87−104).[10]

The spectra were analyzed as a function of denaturant and the absorptions were fit to Gaussians (Figure 2 and Supporting Information) as described previously.[8d] At all denaturant concentrations, the spectra of (d_3)Leu68, (d_8)Lys72, (d_3)Leu94, (d_3)Leu98, and (d_3)Ala101 were well fit by a superposition of the spectrum at zero denaturant (corresponding to the folded protein) and the spectrum at high denaturant, which was identical to that of the free amino acid under the same conditions. In contrast, at intermediate concentrations of denaturant, the spectra of (d_8)Lys73 and (d_3)-Met80 could not be reproduced by a superposition of folded and unfolded spectra, and required an additional Gaussian. The (d_3)-Ala83 spectra were identical at high and low denaturant, but required an additional Gaussian at intermediate concentrations (Supporting Information). The additional Gaussian at each residue reached maximal amplitude at ∼1.2 M GdnHCl and ∼4 M urea, respectively. This signal must correspond to a folding intermediate-(s). An intermediate was also observed with (d_8)Lys79 and GdnHCl, but not urea.

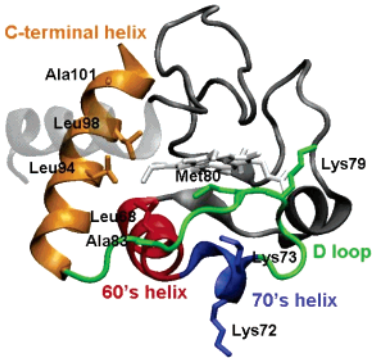

Figure 1. Structure of cyt *c*[10] showing the motifs and residues characterized.

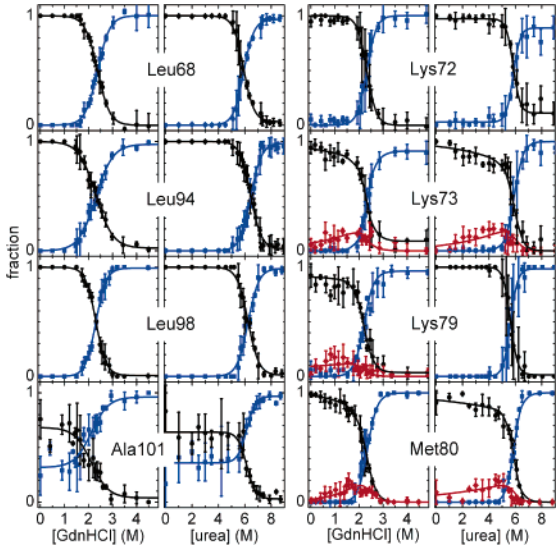

Figure 2. Fractional concentration for folded (black), unfolded (blue), and intermediate (red) cyt *c*, along with fit (lines). Helical fraying at Ala101 results in partial unfolding at zero denaturant.[8d]

The linear extrapolation method, which assumes that ΔG depends linearly on denaturant concentration and has commonly been employed for equilibrium unfolding studies,[11] was used to determine $\Delta G(0)$, the free energy difference between each species in the limit of zero denaturant, *m*, which describes the dependence of ΔG on denaturant concentration, and $c_{1/2}$, the concentration at which both species are present at equal concentrations (Table 1). $\Delta G(0)$ for the folded and unfolded proteins are consistently ∼5 kcal/mol smaller for GdnHCl than for urea. Since the free energy reflects the stability of the same folded state, referenced to the different unfolded states, this suggests that the GdnHCl unfolded state is more stable than the urea unfolded state.

The data reveal correlated behavior of the residues within a structural motif, as well as between the residues of the 60's and

10.1021/ja065179d CCC: $33.50

Table 1. Site-Specific Equilibrium Parameters and Midpoint Concentrations for the Unfolding of Cytochrome *c* in GdnHCl and Urea

			GdnHCl			urea		
equilibrium[a]	motif	residue	ΔG(0) (kcal/mol)	*m* (kcal/M/mol)	$c_{1/2}$ (M)	ΔG(0) (kcal/mol)	*m* (kcal/M/mol)	$c_{1/2}$ (M)
F⇔U	60's helix	Leu68	5.6 ± 0.5	2.3 ± 0.3	2.37 ± 0.05	9.1 ± 0.5	1.5 ± 0.1	5.9 ± 0.1
	70's helix	Lys72	7.1 ± 0.5	3.0 ± 0.3	2.33 ± 0.05	12 ± 2	1.9 ± 0.3	5.8 ± 0.1
		Lys73	8.4 ± 0.6	3.6 ± 0.3	2.29 ± 0.05	11 ± 1	1.9 ± 0.3	5.8 ± 0.2
	D-loop	Lys79	7.6 ± 0.5	3.3 ± 0.3	2.30 ± 0.05	13 ± 1	2.2 ± 0.2	5.6 ± 0.1
		Met80	7.5 ± 0.5	3.2 ± 0.3	2.31 ± 0.05	13 ± 1	2.1 ± 0.2	5.5 ± 0.2
	C-terminal helix	Leu94	4.0 ± 0.2	1.7 ± 0.1	2.31 ± 0.05	9.4 ± 0.7	1.5 ± 0.1	6.4 ± 0.1
		Leu98	6.3 ± 0.3	2.7 ± 0.1	2.31 ± 0.05	9.1 ± 0.7	1.5 ± 0.1	6.2 ± 0.1
		Ala101	4 ± 1	1.7 ± 0.4	2.2 ± 0.1	10 ± 2	1.6 ± 0.3	6.1 ± 0.1
F⇔I	70's helix	Lys73	6.8 ± 0.4	3.3 ± 0.3	2.1 ± 0.2	9 ± 1	1.7 ± 0.2	5.4 ± 0.3
	D-loop	Lys79	5.1 ± 0.4	2.9 ± 0.4	1.8 ± 0.4	*b*	*b*	*b*
		Met80	5.3 ± 0.3	2.6 ± 0.3	2.0 ± 0.2	11 ± 1	2.0 ± 0.2	5.4 ± 0.3
I⇔U	70's helix	Lys73	1.6 ± 0.3	0.3 ± 0.1	4.5 ± 0.5	1.9 ± 0.5	0.2 ± 0.1	10 ± 2
	D-loop	Lys79	1.4 ± 0.4	0.1 ± 0.1	12 ± 3	*b*	*b*	*b*
		Met80	2.2 ± 0.3	0.6 ± 0.1	3.8 ± 0.5	1.6 ± 0.4	0.13 ± 0.05	11 ± 4

[a] F, folded; U, unfolded; I, intermediate state; for fit details see Supporting Information. [b] No intermediate detected.

C-terminal helices and between residues of the 70's helix and D-loop. Residues of the 60's and C-terminal helices have larger $\Delta G(0)$ values and smaller *m* values than residues within 70's helix and D-loop. The apparently cooperative behavior of these folding units (or "foldons"[1c]) suggests that the associated conformational changes involve the entire structural motif, even if not all of the changes are apparent at each residue because of insignificant changes in the environment. The cooperative behavior may be understood from consideration of the protein's structure.[10] The structure reveals that the N-terminal helix packs on both the 60's helix and the C-terminal helix, forming a tightly packed core which apparently unfolds cooperatively. The D-loop and short 70's helix are contiguous, make intimate contacts with the heme cofactor (including the Met80 S−Fe bond as well as packing interactions), and also appear to unfold cooperatively.

The correlated motifs, as well as their relative stabilities, are different from that predicted on the basis of previous experiments.[1c,12] For example, H/D exchange studies suggest that the C-terminal helix unfolds cooperatively with the N-terminal helix (but not the 60's helix) and that this forms the most stable part of the protein. However, these H/D exchange studies were performed at low denaturant, and fluctuations that result in amide exchange are assumed to be similar to those that cause structural changes at higher denaturant. The results of the current study suggest that this is not the case. The current studies show that while the C-terminal and 60's helix do persist to the same or higher denaturant concentrations (i.e. equal or increased $c_{1/2}$ values in GdnHCl and urea, respectively) as the 70's helix and D-loop, this is not the result of greater stability, but rather the result of reduced sensitivity of this cooperative folding unit to denaturant (i.e., lower *m* values). The magnitude of the *m* value is thought to reflect the surface area exposed to solvent during unfolding.[13] Unfolding of the 70's helix and D-loop exposes the large surface area of the heme to solvent, which may explain the large *m* values for this folding unit. Thus, the hydrophobicity of the cofactor plays a central role in folding.

While a variety of techniques have provided evidence of an equilibrium folding intermediate, the intermediate has been difficult to structurally characterize. For example, while NMR and site specific isotope labeling have the potential for high structural resolution, no signals from the intermediate are observable because of the exchange rates that are fast on the NMR chemical shift time scale.[1c] The high time resolution of the IR-based experiments provides an unprecedented direct view of the equilibrium folding intermediate, which appears to have a significantly altered 70's helix and D-loop. The data also reveal that the role of Lys79 appears to be denaturant dependent. While the entire loop that includes Lys79 unfolds cooperatively in both denaturants, the observation of an intermediate at this residue in GdnHCl, but not urea, suggests that the side chain is in a unique environment in the former intermediate, but not the latter. Thus, it seems likely that Lys79 is involved in heme misligation in GdnHCl, as suggested previously, but not in urea.

The incorporation of C−D bonds throughout cyt *c* provides a unique opportunity to characterize the folding process with both high structural and temporal resolution. In addition, it should be possible to time resolve the changes in the C−D spectra after folding is photoinitiated. The combination of equilibrium and time-resolved studies has the potential to fully define the folding mechanism of cyt *c*.

Acknowledgment. This work is based upon research supported by the National Science Foundation under Grant No. 0346969.

Supporting Information Available: Spectral analysis. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) (a) Pletneva, E. V.; Gray, H. B.; Winkler, J. R. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 18397−18402. (b) Jordan, T.; Eads, J. C.; Spiro, T. G. *Protein Sci.* **1995**, *4*, 716−728. (c) Latypov, R. F.; Cheng, H.; Roder, N. A.; Zhang, J.; Roder, H. *J. Mol. Biol.* **2006**, *357*, 1009−1025.

(2) Cinelli, S.; Spinozzi, F.; Itri, R.; Finet, S.; Carsughi, F.; Onori, G.; Mariani, P. *Biophys. J.* **2001**, *81*, 3522−3533.

(3) Lyubovitsky, J. G.; Gray, H. B.; Winkler, J. R. *J. Am. Chem. Soc.* **2002**, *124*, 14840−14841.

(4) Thomas, Y. G.; Goldbeck, R. A.; Kligar, D. S. *Biopolym. (Biospectros.)* **2000**, *57*, 29−36.

(5) Bai, Y. W.; Sosnick, T. R.; Mayne, L.; Englander, S. W. *Science* **1995**, *269*, 192−197.

(6) Filosa, A.; Wang, Y.; Ismail, A. A.; English, A. M. *Biochemistry* **2001**, *40*, 8256−8263.

(7) Russell, B. S.; Melenkivitz, R.; Bren, K. L. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *15*, 8312−8317.

(8) (a) Chin, J. K.; Jimenez, R.; Romesberg, F. E. *J. Am. Chem. Soc.* **2001**, *123*, 2426−2427. (b) Chin, J. K.; Jimenez, R.; Romesberg, F. E. *J. Am. Chem. Soc.* **2002**, *124*, 1846−1847. (c) Sagle, L. B.; Zimmermann, J.; Dawson, P. E.; Romesberg, F. E. *J. Am. Chem. Soc.* **2004**, *126*, 3384−3385. (d) Sagle, L. B.; Zimmermann, J.; Matsuda, S.; Dawson, P. E.; Romesberg, F. E. *J. Am. Chem. Soc.* **2006**, *128*, 7809−7915.

(9) Residues 70−75 appear to form a helix in the NMR structure and a loop in the crystal structure (see ref 11).

(10) Bushnell, G. W.; Louie, G. V.; Brayer, G. D. *J. Mol. Biol.* **1990**, *214*, 585−595.

(11) Greene, R. F., Jr.; Pace, C. N. *J. Biol. Chem.* **1974**, *249*, 5388−5393.

(12) Milne, J. S.; Mayne, L.; Roder, H.; Wand, A. J.; Englander, S. W. *Protein Sci.* **1998**, *7*, 739−745.

(13) Tanford, C. *J. Am. Chem. Soc.* **1964**, *86*, 2050−2059.

JA065179D